

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:27:50 : Search time 11.9435 Seconds
(without alignments)
979.302 Million cell updates/sec

Title: US-09 930-026-2

Perfect score: 151

Sequence: 1 MAUQCM:SSOKA:MLHLSA.....AMAPQKP:HSQWGNHSSC 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum Match 98

Maximum Match 150

Listing first 135 summaries

Database: SWISSPROT_40.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	1241	62.1	236	1	UBC3_RABIT	Q29503 oryctolagus
2	1010	66.8	236	1	UBC3_HUMAN	P49427 homo sapien
3	460.5	30.5	170	1	UBC6_HUMAN	Q99462 homo sapien
4	452	29.9	167	1	UBC7_ARATH	P42747 arabidopsis
5	432.5	28.6	225	1	UBC3_YEAST	P14682 saccharomyc
6	428.5	28.4	164	1	UBC7_CAEEL	P34477 caenorhabdi
7	423.5	28.2	213	1	UBC2_ASFPM2	P27949 african swi
8	411.5	27.2	215	1	UBC3_HUMAN	P25868 triticum ae
9	404.5	26.8	168	1	UBC7_HUMAN	P56554 homo sapien
10	380.5	25.2	165	1	UBC7_HUMAN	Q00102 schizosacch
11	367	24.3	166	1	UBC7_SCHPO	Q02159 saccharomyc
12	356.5	23.7	165	1	UBC2_YEAST	P06104 saccharomyc
13	310	20.5	172	1	UBC2_CANAL	O74201 candida alb
14	308	20.4	179	1	UBC1_CAEEL	P52478 caenorhabdi
15	301.5	20.0	192	1	UBC2_CANAL	P23567 homo sapien
16	295	19.5	152	1	UBC2_HUMAN	P49459 homo sapien
17	286	18.9	152	1	UBC2_HUMAN	P23566 schizosacch
18	284	18.6	151	1	UBC2_SCHPO	P52493 neurospora
19	281	18.6	151	1	UBC2_NEUR	P42745 arabidopsis
20	280	18.5	152	1	UBC2_ARATH	P35130 medicago sa
21	278	18.4	152	1	UBC2_MEDSA	P25153 drosophila
22	277	18.3	151	1	UBC6_DROME	P25866 triticum ae
23	276	18.3	152	1	UBC1_ARATH	O00103 schizosacch
24	276	18.3	152	1	UBC2_WHEAT	P42746 arabidopsis
25	267.5	17.7	176	1	UBC2_SCHPO	P25153 drosophila
26	266	17.6	150	1	UBC3_ARATH	P25866 triticum ae
27	265.5	17.6	156	1	UBC4_YEAST	O00103 schizosacch
28	264.5	17.3	148	1	UBC3_SCHPO	P42746 arabidopsis
29	261	17.3	148	1	UBC3_SCHPO	Q95044 spissula sol
30	260.5	17.2	157	1	UBC3_SCHPO	P35134 arabidopsis
31	254.5	16.8	179	1	UBC3_HUMAN	P40984 schizosacch
32	252	16.7	248	1	UBC4_YEAST	O00762 homo sapien
33	252	16.7	148	1	UBC9_ARATH	P35132 arabidopsis

P51668	homo sapien
P51131	arabidopsis
P51133	arabidopsis
P66516	xenopus lae
P15731	saccharomyc
P51228	drosophila
P21734	saccharomyc
P25867	drosophila
P43102	candida alb
P50530	homo sapien
P51229	caenorhabdi
P51966	homo sapien
P15732	saccharomyc
P46595	schizosacch
P28263	saccharomyc
P47986	homo sapien
P56617	arabidopsis
P70711	rattus norv
Q16781	homo sapien
P52490	saccharomyc
P51669	homo sapien
O09181	mesocricetu
Q95111	schizosacch
P50623	saccharomyc
P42750	arabidopsis
P42749	arabidopsis
P51965	homo sapien
P52482	mus musculu
P52483	mus musculu
P42748	arabidopsis
Q16763	homo sapien
P52485	drosophila
P27924	homo sapien
G14933	homo sapien
P42743	arabidopsis
P37286	homo sapien
P16577	triticum ae
P49428	pichia past
O60015	pichia anqu
P52487	drosophila
P52486	drosophila
Q96909	homo sapien
P52491	saccharomyc
P29340	saccharomyc
P32296	saccharomyc
P52484	caenorhabdi
G11076	caenorhabdi
P11231	rana esculu
Q10175	schizosacch
P14211	mus musculu
P32255	saccharomyc
P18418	rattus norv
P15253	oryctolagus
P53165	saccharomyc
P07746	oncorhynch
P27692	saccharomyc
P19204	gallus gall
P35947	drosophila
P07155	mus musculu
P31119	daucus caro
P22414	candida tro
P06429	homo sapien
O09185	mus musculu
P38996	saccharomyc
P40021	saccharomyc
P12685	saccharomyc
P10103	bos taurus
P25979	xenopus lae
P15208	mus musculu
P06213	homo sapien
P5127	rattus norv
P12682	sus scrofa

107 53 6.2 6.02 1 SPT8_YEAST
108 93 6.2 7.05 1 PIXH_MOUSE
109 92.5 6.1 1032 1 KINN_HUMAN
110 92.5 6.1 2493 1 CYAA_USMA
111 91.5 6.1 211 1 HMTX_HUMAN
112 91.5 6.1 732 1 YMBK_YEAST
113 91.5 6.1 913 1 IF38_HUMAN
114 91.5 6.1 1027 1 KINN_MOUSE
115 91 6.0 294 1 NPM_HUMAN
116 91 6.0 543 1 C314_DROME
117 91 6.0 700 1 CH60_PLAFG
118 91 6.0 2452 1 RBL1_PLAFD
119 91 6.0 3562 1 PCVC_CHICK
120 90.5 6.0 292 1 NPM_MOUSE
121 92.5 6.0 366 1 IZAA_DROME
122 90.5 6.0 417 1 CRTIC_HUMAN
123 90 6.0 1094 1 YH00_YEAST
124 89.5 5.9 470 1 ESCA_DROME
125 89 5.9 415 1 CAQC_MOUSE
126 89 5.9 840 1 YUTL_YEAST
127 89 5.9 1132 1 YKKS_YEAST
128 89 5.9 1276 1 PRDG_HUMAN
129 89 5.9 1553 1 TP2A_CHICK
130 88.5 5.9 233 1 CRXC_NITEU
131 88.5 5.9 416 1 CRIC_BETVU
132 88 5.8 401 1 HB9_HUMAN
133 88 5.8 421 1 CRTIC_PRUAR
134 88 5.8 459 1 T.G_MYCPU
135 88 5.8 638 1 GHR_RAT

ALIGNMENTS

RESULT 1
JBC3_RABIT
ID UBC3_RABIT STANDARD: PRT: 238 AA.
AC Q29503;
DT 03-NOV-1997 (Rel. 35, Created)
DT 03-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-32 kDa complementing (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (E2-CDC34).
SN CUC34.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE: Heart;
RX MEDLINE=97236448; PubMed=9116038;
RA Sun B.G., Jayaseelan K., Chung M.C., Tan T.W., Chock P.B., Ieo T.S.;
RT "Cloning, characterization and expression of a cDNA clone encoding
rabbit ubiquitin-conjugating enzyme, E2(32k).";
RL Biochim. Biophys. Acta 1351:231-238(1997).
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS.
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: ubiquitin conjugation; second step.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC STRONGEST, TO YEAST UBC3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb.sib.ch).

CC EMBL: U58652; AAB02656.1; ALT_INIT.
DR HSP: Q02159; 2UCZ
DR InterPro: IPR000508; UHQ_conjugat.
DR Pfam: PF00179; UO_con; 1.
DR ProDom: PD000461; UHQ_conjugat; 1.
DR SMART: SM00212; UBCr; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGT_2; 1.
KW ubl conjugation pathway; Ligase; Multigene family.
FT BINDING 93 93 UBIQUITIN (BY SIMILARITY).
FT DOMAIN 200 238 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 238 AA: 27166 MW: E896CF0116A56308 CRC64;
Query Match 82.18; Score 1241; Db 1; Length 238;
Best Local Similarity 99.86; Pred. No. 3.8e-92;
Matches 233; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAQOMTSQKALMLKLSQEEPPVEGFRITLVDESILYNNVAIEGLPNTLYEGGYEKA 60
DB 1 MAQOMTSQKALMLKLSQEEPPVEGFRITLVDESILYNNVAIEGLPNTLYEGGYEKA 60
QY 61 HIKFIDIPYSPPTFRFLTKMHPNIYENGWCISILHPPVDDPQSGELPSERWNPUNV 120
DB 61 HIKFIDIPYSPPTFRFLTKMHPNIYENGWCISILHPPVDDPQSGELPSERWNPUNV 120
QY 121 RTLLSVLSILNPNIFSPANVDASVYFKWDSKGDKYAKTIKQVSATKAAEKDG 180
DB 121 RTLLSVLSILNPNIFSPANVDASVYFKWDSKGDKYAKTIKQVSATKAAEKDG 180
QY 181 VKVPTTLAEYCIKTKVPSNDSSDLLYDDDDIDDEDEDEEDADCYDDDDSG 234
DB 181 VKVPTTLAEYCIKTKVPSNDSSDLLYDDDDIDDEDEDEEDADCYDDDDSG 234

RESULT 2
UBC3_HUMAN
ID UBC3_HUMAN STANDARD: PRT: 236 AA.
AC P49427;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-32 kDa complementing (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (E2-CDC34).
GN CDC34.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94088425; PubMed=8248134;
RA Plon S.F., Leppig K.A., Do H.N., Groudine M.;
RT "Cloning of the human homolog of the CDC34 cell cycle gene by
RT complementation in yeast."
RL Proc. Natl. Acad. Sci. U.S.A. 90:10484-10488(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE: Brain, and Lung;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS.
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC STRONGEST, TO YEAST UBC3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

SEQUENCE FROM N.A.
SPECIES-Mouse; STRAIN-C57BL/6J; TISSUE-Head;

SPECIES-Mouse: STRAIN-C57BL/6J; TISSUE-Head:
MEDLINE-21085660; PubMed-11217851.
Kawai J., Shinnagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Kouno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishio K., Kiyosawa H., Kondo S., Yamakawa I.,
Saito T., Okazaki Y., Gijobori T., Bono H., Kasukawa I., Saito K.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Koebnick M.,

Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L. M., Staubli F., Suzuki K., Tomita M., Wagner U., Washio T., Sakai K., Okido I., Furuno M., Aono H., Baldarelli R., Harsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Gariboldi M. F., Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustinchac S., Hill D., Hofmann M., Humé D. A., Kamiya M., Lee N. H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordore P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Serebriy I., Shih C., Shih C.,

Suzuki H., Toyooka K., Wang K.H., Wetz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;

Functional annotation of a full-length mouse cDNA collection. Nature 409:685-690(2001).

[4] SEQUENCE FROM N. A.

SPECIES-Rat;
MF01 INF-98262616; Pubmed 10355553

LIBRARY 0202010, FUDMEI-10329603;
Lic H., Wing S.S.;

*Identification of rabbit reticulocyte E217K as a UIC7 homologue and functional characterization of its core domain loop.

J. Biol. Chem. 274:14685-14691(1999).

PROTEIN. CHANGES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS. MAY BE INVOLVED IN DEGRADATION OF MUSCLE-SPECIFIC PROTEINS.

-1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +

-1- PATHWAY: Ubiquitin conjugation; second step

- TISSUE SPECIFICITY: WIDELY EXPRESSED, HIGHER IN SKELETAL MUSCLE.
- MISCELLANEOUS: A CYSTEINE PROTEINOLYTIC ENZYME.

UBIQUITIN-THIOLESTER FORMATION.

1. SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY. STRONGEST, TO C.ELEGANS UBC7.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute.

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute.

use by non-profit institutions as long as its content is in no way modified and that the European Bioinformatics Institute. There are no restrictions on its

entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>)

-----or send an email to license@lsb-sib.ch).

EMBL: D78514; BAA11410.1; -.
EMBL: BC002775. AAC00275.1. -

EMBL; AK013902; BAH29048.1; -

EMBL; AF099093; AAC69605.1; -;
HSSP: Q02159; 2UGZ.

Genew; HGNC:12482; UBE2G1.
MIM: 601569; -

InterPro; IPR000608; UHQ_conjugat.
 Pfam: PF00170. 10 cons.

Problem; PD000461; URQ_conjugat; 1.

SMART; SM00212; UBCC; 1.
PROSITE; PS00183; UBIQUITIN CONJUGAT 1.]

PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.

sex conjugation pathway, ligase; Multigene family.
BINDING 90 90 UBIQUITIN (BY SIMILARITY).

SEQUENCE 170 AA; 19509 MW; 36B61766D995R332 CRC64;

ry Match
30.5%; Score 460.5; DB 1; Length 170;

Host local Similarity 52.4%; Prod. No. 4.70-30;
Matches 97; Conservative 28; Mismatches 44; Indels 7; Gaps 3;

QY 6 MTSSKALML--ELKSQGEPEVGFITLIVDSLYNWEVAIFGLPNTLYEGGYFAHAK 63
DB : MTELASALLRQIAELINKNPVGFSGAIDNDNLYRWEVILISPPDLYEGGVFAHIT 60
QY 64 PFDYPSPTFRPLTKMMHPNIYFGDVCISILHPVDPOSGELPSFRKNWNTONVETI 123
DB : MTELASALLRQIAELINKNPVGFSGAIDNDNLYRWEVILISPPDLYEGGVFAHIT 120

QY 124 LLSVLSLLNPNFTSPANVDASVMPKRRKSDKGD--KEYAFIRK 167
DB : MTSV:SMIADPNCISPNVZAA---KENREDNGEKKKVARCVK 163

RESULT 4
UBC7_ARATH
ID UBC7_ARATH STANDARD: PRT: 167 AA.
AC P42747;
DI 01-NOV-1995 (Rel. 32, Created)
DI 01-NOV-1995 (Rel. 32, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-19 kDa (EC 6.3.2.19) (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (TAT029).
GN UBC7 OR UBC14 OR AT3G55380 OR T22B16.40.
CS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; Curosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RX MEDLINE-94359481; PubMed-8078482;
RA Genschik P., Durr A., Fleck J.;
RT "Differentia; expression of several E2-type ubiquitin carrier protein genes at different developmental stages in Arabidopsis thaliana and Nicotiana sylvestris";
RL Mol. Gen. Genet. 244:548-556(1994).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RX MEDLINE-96218124; PubMed-8647807;
RA van Nocker S., Walker J.M., Vierstra R.D.;
RT "The Arabidopsis thaliana UBC7/13/14 genes encode a family of multi-ubiquitin chain-forming E2 enzymes";
RL J. Biol. Chem. 273:12150-12158(1998).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RX MEDLINE-21016720; PubMed-11130713;
RA Salanoubat M., Iemcke K., Rieger M., Ausorge W., Unseld M., Fartmann B., Hutter G., Blocker H., Perez-Alonso M., Obermaier H., Delser M., Bally M., Grivell I.A., Mache K., Paigdomonech P., de Simone V., Choise N., Artiguenave F., Robert C., Brottier P., Winkler P., Calolico L., Weissbach J., Saurin W., Quetier F., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V., Wurmbach E., Drzonek H., Erlie H., Jordan N., Bangert S., Wiedemann K., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G., Verzi A., d'Angelo M., Pallavicini A., Toppi S., Simonati R., Conrad A., Hornischer K., Kauer G., Lochner T.-H., Nordstok G., Reichelt J., Scharfe M., Schron O., Barges M., Terol J., Clement J., Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D., Cooke K., Landie M., Berger-Llauro C., Purnelle B., Masuy D., de Haar M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta R., Monfort A., Agrillon A., Flores M., Liguori R., Vitale D., Mannbapt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W., Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins C., Rooney T., Rizzo M., Wallis A., Utterback T., Fajli C.Y., Shea T.P., Creusy T.H., Haas B., Maiti K., Wu D., Peterson J., Van Aken S., Pat G., Millischer J., Sellers P., Gill J.E., Feldblum T.V., Preuss D., Lin X., Nicotian W.C., Salzberg S.L., White O., Venter J.C.,

Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu F., Sasamoto S., Kimura T., Iidesawa K., Kawashima K., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis thaliana";
RL Nature 408:820-822(2000).
RN [4]
SEQUENCE OF 1-67 FROM N.A.
RC STRAIN-cv. Columbia;
RA Borthomieu P., Guerrier D., Giraudat J.;
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP + diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOLESTER FORMATION.
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X72625; CAAS1200.1;
DR EMBL: U3759; AAC49323.1;
DR EMBL: AL132975; CAB75896.1;
DR EMBL: Z18513; CAA79212.1;
DR HSP: Q62159; 2UCZ.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UQ_con: 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBQC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubiquitin-conjugating pathway; Ligase; Multi-subunit family.
FT BINDING 90 90 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 167 AA; 18728 MW; 15053C16A62C08BC CRC64;
Query Match 29.9%; Score 452; DB 1; Length 167;
Best Local Similarity 50.9%; Pred. No. 2.2e-29;
Matches 85; Conservative 33; Mismatches 37; Indels 12; Gaps 3;
QY 6 MTSSKALML--ELKSQGEPEVGFITLIVDSLYNWEVAIFGLPNTLYEGGYFAHAK 63
DB : MANNQASILLQKLDCKKPKVDGFSAGLVDEKKNVFSVSVSGPPDLYEGGFNAIMS 60
QY 64 PFDYPSPTFRPLTKMMHPNIYFGDVCISILHPVDPOSGELPSFRKNWNTONVETI 123
DB : PPNYPSPTVTFTSEMMHPNVYSDGKVCISILHPVDPOGPHCYELASERKTPVHTVESI 120
QY 124 LLSVLSLLNPNFTSPANVDASVMPKRRKSDKGDKEYAFIRK 170
DB : VLSIISLISLGNPDESANVEA---KRWONRAE-----PRKVS 157
RESULT 5
UBC3_YEAST
ID UBC3_YEAST STANDARD: PRT: 295 AA.
AC P14682;
DI 01-APR-1990 (Rel. 14, Created)
DI 01-APR-1990 (Rel. 14, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-34 kDa (EC 6.3.2.19) (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (Cell division control protein 34).
GN UBC3 OR CDC34 OR DNA6 OR YDR054C OR YD9609.08C OR D4211.

DR ProDom: P0000461; UDO conjugat. 1.
 DR SMART: SM00212; UBC: 1.
 DR PROSITE: PS00183; UBIQUITIN-CONJUGAT_1; 1.
 DR PROSITE: PS00127; UBIQUITIN-CONJUGAT_2; 1.
 KW Ubl conjugation pathway: ligase; Multigene family.
 FT BINDING 88 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 164 AA: 18938 MW: 625800DF8F8D958 CRC64;
 Query Match 28.4%; Score 428.5; DB 1; Length 164;
 Best Local Similarity 48.8%; Pred. No. 1.6e-27;
 Matches 83; Conservative 25; Mismatches 56; Indels 9; Gaps 3;
 QY 6 MTSSKALMLLEIKSQEPVEGRITLVDSDLYNWEVAIFGLPNTLYEGGYFKAHKEP 65
 Db 1 MGSSLLKKQLADRRVVDGASAGLVNDNLYKWEVLVCPDIIILYGGFFKATLDP 60
 QY 65 IDYPSPTFFRLKMMHNPNIYSDVCSISIHPPVDDPQSGE:PSRNMNPTONVITILL 125
 Db 61 RDPYQPKRMKAFISRIWHPNDKGNVCISILHDPGDDKWKGYEPEERNLPVITVETILL 120
 QY 126 SVISLLEPNPFIESPANVDASVWFKKRSKDKKDYAEIRKQVSAKAEAK 178
 Db 121 SVISLITOPNFSPANVDAAMKOR-----ENYAEKKKKVACQVRRSQEE 164
 RESULT 7
 URC_ASPM2
 ID URC_ASPM2 STANDARD; PRT: 213 AA.
 AC P25669;
 DT 31-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubl-ubiquitin-conjugating enzyme E2-21 kDa (EC 6.3.2.19)
 DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
 GN ASPV-UBC.
 OS African swine fever virus (isolate Malawi Lil 20/1) (ASFV).
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarvirus.
 OX NCBI_TaxID:10500;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE:92255177; PubMed:1110934;
 RA Bingamp P.M., Arnold J.F., Mayer R.J., Dixon L.K.;
 RT "A ubiquitin conjugating enzyme encoded by African swine fever virus".
 RC EMBO J. 11:361-365 (1992).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS.
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP + diphosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOLESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC EMBL: X62449; CA444305.1;
 DR EMBL: X71982; CA450851.1;
 DR PIR: S19158; S19158.
 DR HSRP: Q02159; 2002.
 DR InterPro: IPR000608; URC_conjugat.
 DR Pfam: PF00179; URC_con.1.
 DR ProDom: PD000461; URC_conjugat.1.
 DR SMART: SM00212; URC: 1.
 DR PROSITE: PS00183; UBIQUITIN-CONJUGAT_1; 1.
 DR PROSITE: PS00127; UBIQUITIN-CONJUGAT_2; 1.
 KW Ubl conjugation pathway: ligase.

FT BINDING 85 UBIQUITIN (BY SIMILARITY).
 FI DOMAIN 183 ASP/GLU-RICH (ACIDIC).
 SQ SEQUENCE 213 AA: 24468 MW: F9671BC7385D6DCE CRC64;
 Query Match 28.2%; Score 425.5; DB 1; Length 213;
 Best Local Similarity 42.8%; Pred. No. 3.9e-27;
 Matches 95; Conservative 33; Mismatches 77; Indels 17; Gaps 5;
 QY 13 IMLELAKSQEPVEGRITLVDSDLYNWEVAIFGLPNTLYEGGYFKAHKEPDIYPSP 72
 Db 6 LLAFYKNI.LVNSEHFKLS-VNEDNLTEWDDVLKPPDPIYEGGLFKAK:VEPKPYEP 64
 QY 73 PTERETKMMHNPNIYSDVCSISIHPPVDDPQSGELPSEKRNPTONVITILLSVISLN 132
 Db 65 PRUFTSEMMHNPNIYSDKLCISILH-----GUNEQGMTHSPAQKIDTLVLSISLN 119
 QY 133 EPWTFSPANVDASVWFKR-----RDSKCKKAEAEIRKQVSAKAEAKGVKVTTLAE 189
 Db 120 EPNPDSPANVDAKSYRKYLYKEDLESYPMEVKTKYKSLDECSAEDIEYKKNPVNV-- 177
 QY 190 YCLTKVPSNDNSLLYDDLYDDIDDEDEDEEDADCYDD 231
 Db 178 ----LPVPSDDYEDHEDGTYLLTYDDEDEDEDEE--MDDF 213
 RESULT 8
 URC_ASPB7
 ID URC_ASPB7 STANDARD; PRT: 215 AA.
 AC P27949;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-21 kDa (EC 6.3.2.19)
 DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
 GN E215L.
 OS African swine fever virus (strain BA71V) (ASFV).
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarvirus.
 OX NCBI_TaxID:10498;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE:92087485; PubMed:1309282;
 RA Rodriguez J.M., Salas M.L., Vinuela E.;
 RT "Genes homologous to ubiquitin-conjugating proteins and eukaryotic transcription factor SII in African swine fever virus".
 RL Virology 186:40-52 (1992).
 RN 12;
 RP COMPLETE GENOME.
 RA Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C., Rodriguez J.F., Vinuela E.;
 RT "Analysis of the complete nucleotide sequence of African swine fever virus".
 RL Virology 208:249-278 (1995).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS.
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP + diphosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- DEVELOPMENTAL STAGE: MAINLY LATE IN INFECTION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOLESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC EMBL: M7121; AAA42704.1;
 DR EMBL: U18466; AAA65370.1;
 DR PIR: F39448; UGXFAS.

```

CC
DR EMBL: M74077; ; NOT_ANNOTATED_CDS.
DR PIR: A41547; A41547.
DR HSP: Q02159; ZUC7.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UQ_con: 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBQC: 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
DR Ubl conjugation pathway; Ligase: Multigene family.
KW BINDING 92 92 UBIQUITIN (BY SIMILARITY).
FT BINDING 92 92 UBIQUITIN (BY SIMILARITY).
SQ
SEQUENCE 168 AA; 18897 MW; 6ED5127DCAB415E2 CRC64.

Query Match 26.88; Score 404.5; DH 1; Length 168;
Best Local Similarity 50.78; Pred. No. 1.4e-25;
Matches 74; Conservative 28; Mismatches 37; Indels 7; Gaps

QY 24 PVGEFRITLVDESILYNNVEAIGLPTNLYEGYFKAIKFPIDYSPPTFRELTHMHII 83
DQ 23 PSMGFQLGVDOSNVFQVITIGPPEIHYDGGYENATMGSPNPNSPTPTRETSMMH 82
QY 84 PNLYENGDCVCSILHPVPDDPSGELSPERNPQNVTILLSVISILNEPNTSPANVD 143
DQ 83 PNLYPDGRGVCSILHPGDDPFGYELASERMPVITVESIVLSIISMSSNDESPANIE 141
QY 144 ASVMEFKKRWDSKQKQKPYAELLRKQV 169
DQ 142 AA---KDWRE---KODEFKKKVRRAV 161

RESULT 10
UBCJ HUMAN
ID UBCJ3 HUMAN STANDARD; PRT: 165 AA.
AC P56554;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2 G2 [EC 6.3.2.19] (Ubiquitin-protein
DE ligase G2) (Ubiquitin carrier protein G2).
GN UBC2G2.
GC Homo sapiens (Human), and
GC Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=98360100; PubMed=9693041;
RA Katsanis N., Fisher F.M.C.;
RT *Identification, expression, and chromosomal localization of ubiquitin-
RT conjugating enzyme 7 (UBC2G2), a human homologue of the Saccharomyces
RT cerevisiae UBC7 gene.;
RN [2]
RN Genomics 51:128-131(1998).
RC SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Tokoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakai Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Katsuyuki K., Asakawa S.,
RA Shintani A., Sakaki T., Nagamine K., Miyasawa S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P.,
RA Schorfme M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riessemann L., Dayand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Zietz D., Francis F.,
RA Schrach H., Reinhardt K., Faspo M.-L.;
RT *The DNA sequence of human chromosome 21.*;
RN Nature 405:311-319(2000).

```

```

RN  [3]
RP  SEQUENCE FROM N.A.
RC  SPECIES=Human; TISSUE=Lung;
RA  Strausberg R.;
RL  Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN  [4]
RP  SEQUENCE FROM N.A.
RC  SPECIES=Mouse; STRAIN=C57BL/6J; TISSUE=Fetal;
RA  MEDLINE=21238254; PubMed 11278356;
RA  Tiwari S., Weissman A.M.;
RT  "Endoplasmic reticulum (ER)-associated degradation of T cell receptor
K2 subunits. Involvement of EF-associated ubiquitin-conjugating enzymes
RT (E2s).";
RL  J. Biol. Chem. 276:16193-16200(2001).
RN  [5]
RP  SEQUENCE FROM N.A.
RC  SPECIES=Mouse;
RA  Strausberg R.;
RL  Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC  -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC  OTHER PROTEINS.
CC  -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
CC  diphosphate + protein N-ubiquityl-lysine.
CC  -1- PATHWAY: Ubiquitin conjugation; second step.
CC  -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC  UBIQUITIN-THIOLESTER FORMATION.
CC  -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC  STRONGEST, TO C-TERMINUS UBC7.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
ENBL; AF032456; AAC32312.1;
EMBL; AL163300; CAB90551.1;
EMBL; H0501738; AAH01738.1;
EMBL; H0509355; AAH09351.1;
EMBL; BC011569; AAH11569.1;
EMBL; AF296657; AAK52608.1;
EMBL; BC010321; AAH10321.1;
HSP; Q02159; 2UC2.
Gene; H0501738; JBR232.
MD; MG1343188; Ube2a2.
InterPro: IPR000608; Ubc7_conjugat.
Pfam: PF00179; Ubc7_conj_1.
ProDom: PD000461; Ubc7_conjugat; 1.
SMART: SM0212; Ubc7; 1.
PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
Ubl conjugation pathway; Lysase; Multigene family.
BINDING 89 89 UBIQUITIN (BY SIMILARITY).
FT BINDING 12 12 E > V (IN REF. 1).
FT CONFID 191 107 MOTESSA -> HC-REQ (IN REF. 1).
SQ SEQUENCE 165 AA; 18+66 MW; 74D6C732A79575E3 CRC64;
Query Match:
Best local similarity 55.3%; Score 380.5; DB 1; Length 165;
Matches 80; Conservative 18; Mismatches 54; Indels 7; Gaps 2;
Qy 11 KAIIMELKSIQEEVPEGRPTLVIESDLYNVEVAIFGLNPLIFGCGYFKAHKFPIDY 70
2b 7 KRLMAEYKULTINPEGSVAIPKMEENEFEEALIMGPEDYCFEGYVFAILSFPLDPL 66
Qy 71 SPPTFRITKMKHNDYNGVNCVLSILHPVDHPQSCHEPSEKRNPTQNVRTTLLSVTS 130
Db 67 SPPEKMFRTFEMFHRNIYDGRVGLSLHARGDQPMGYSSAERNSPQSVKILISVSM 126
Qy 131 INEPTFTSPANDVASVMEKRWKRSKUKKEYAEIIRKQV 169

```

```

Db 127 LAEPNDESGANVDASKM---WRD---DREQFYKIAKQI 158
RESULT 11
ID UBC7_SCHPO STANDARD: PRT; 166 AA.
AC 000102; G9HDP3;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-18 kDa (EC 6.3.2.19) (Ubiquitin-
DE protein ligase) (Ubiquitin carrier protein).
GN UBCP3 OR SPBP16F5.04.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=97295689; PubMed=9154838;
RA Osaka F., Sano H., Sano T., Yamao F.;
RT "A ubiquitin-conjugating enzyme in fission yeast that is essential
RT for the onset of anaphase in mitosis.";
RL Mol. Cell. Biol. 17:3388-3397(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell I., Fraser A.,
RA Gentles M., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Quail M.A., Rabinowitsch E.,
RA Skellon J., Simmonds M., Squares R., Stevens K., Sharp S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren I., Whitehead S.,
RA Woodward J., Wolckaert G., Aert R., Robben J., Grymoprez B.,
RA Welljens I., Vanstraelen E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Horzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl I.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadiou E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
CC diphosphate + protein N-ubiquityl-lysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC EMBL: D85544; HAA20473.1;
CC EMBL: A1441663; CAC08543.3;
CC HSSP: Q02157; 20C2.
CC InterPro: IPR000608; Ubq_conjugat.
CC Pfam: PF00179; Ubq_con; 1.
CC PRODOM: P000461; Ubq_conjugat; 1.
CC SMART: SM00212; Ubq; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS01227; UBIQUITIN_CONJUGAT_2; 1.
KW Ub1 conjugation pathway; Ligase; Multiye gene family.
FT BINDING 90 90 UBIQUITIN (BY SIMILARITY).
FT COMPLECT 6 6 A -> P (N REF. 1).
SQ SEQUENCE 166 AA; 18720 MW; C376AC652EF444EE CRC64;

Query Match 24.3%; Score 367; DB 1; Length 166;
Best Local Similarity 45.1%; Pred. No. 1.3e-22;
Matches 71; Conservative 25; Mismatches 58; Indels 6; Gaps 2;

QY 11 KALMLKLSIQEPVSGFRHILVDESOLYNNWEVA:EGIPNTLYEGGYFKAHIKPPTIDY 70
DB 6 KRLMKYKELTENGPGITAGSNEDEFTWCCLIQGPDGTFEGGLYPATLKFPDSYPL 67
QY 71 SPTFRFLKMWHPNTYENGWVCSILIPPVDPQSGCELPSEMMNPQNVRIILLWSVIL 130
DB 68 GPPILAFEGEFPHPVYKDGIVCSILHACDOPNNYESSERSPVOSVKEKILLWSNM 127
QY 131 LNEPNTFSPANDVSMFKNRDSSGKDKKEYAEIRKQVSAT 172
DB 128 LAEPNDESGANIDACKM---WRDR---PYCNVVRRIAKRT 163

RESULT 12
UBC2_YEAST
ID UBC2_YEAST STANDARD: PRT; 165 AA.
AC Q02157;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Ubiquitin-conjugating enzyme E2-18 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
GN Ubc7 OR Q18 OR YMR022W OR YH9711.12.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB320;
RX MEDLINE=93003327; PubMed 1327148;
RA Vassal A., Boulet A., Decoster F., Faye G.;
RI "Q18, a novel ubiquitin-conjugating enzyme in Saccharomycetes cerevisiae.";
RJ Biochim. Biophys. Acta 1132:211-213(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93149278; PubMed=8381213;
RA Jungmann C., Reins H.-A., Schober C., Jentsch S.;
RI "Resistance to cadmium mediated by ubiquitin-dependent proteolysis.";
RJ Nature 361:369 371(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S286c / AB972;
RX Lye G., Churcher C.M., Barrell B.G., Rajadream M.A., Walsh S.V.;
RI Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.93 ANGSTROMS).
RX MEDLINE=97200716; PubMed=9048545;
RA Cook W.J., Marlin P.D., Edwards B.F.P., Yamazaki R.K., Chau V.;
RI "Crystal structure of a class I ubiquitin conjugating enzyme (Ubc7) from Saccharomyces cerevisiae at 2.9-A resolution.";
RJ Biochemistry 36:1621-1627(1997).
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER

```

```

CC PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED AND
CC ABNORMAL PROTEINS. INVOLVED IN RESISTANCE TO CADMIUM
CC POISONING.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine ~ AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- INDUCTION: BY CADMIUM.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X66829; CAA47302.1;
CC EMBL: X69100; CAA48846.1;
CC EMBL: X49211; CAA89125.1;
CC PIR: S28951; S28951
CC PIR: S29338; S29338
CC PIR: S29741; S29741
CC PIR: 20C2; 18-MAR-98.
CC SGD: S0004624; Q18.
CC InterPro: IPR003608; Ubq_conjugat.
CC Pfam: PF00179; Ubq_con; 1.
CC PRODOM: P000461; Ubq_conjugat; 1.
CC SMART: SM00212; Ubq; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS01227; UBIQUITIN_CONJUGAT_2; 1.
KW Ub1 conjugation pathway; Ligase; Multiye gene family; 3D-structure;
KW Cadmium.
FT BINDING 89 89 UBIQUITIN.
SQ SEQUENCE 165 AA; 18520 MW; D3D297847DBB462D CRC64;

Query Match 23.7%; Score 358.5; DB 1; Length 165;
Best Local Similarity 47.3%; Pred. No. 6.3e-22;
Matches 70; Conservative 26; Mismatches 49; Indels 3; Gaps 1;

QY 8 SSQALMLKLSIQEPVSGFRHILVDESOLYNNWEVA:EGIPNTLYEGGYFKAHIKPPTID 67
DB 4 TAQRRLKELQQLKIDSPGIVAGPKSENNIFTWCCLIQGPDTPYADGVENAKLEFPKD 63
QY 66 YPSPTFRFLKMWHPNTYENGWVCSILIPPVDPQSGCELPSEMMNPQNVRIILLWSV 127
DB 64 YPLSPPKLIFTPSILHPNLYPNGEVCSILHSPCDOPNNYELAEERWSPVQSVFKILLSV 123
QY 128 ISLLNEPNTFSPANDVSMFKNRDSSK 155
DB 124 MSMLSEPNIESGANIDACKL---WRDNR 148

RESULT 13
UBC2_YEAST
ID UBC2_YEAST STANDARD: PRT; 172 AA.
AC P06104;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-20 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
GN UBC2 OR RAU6 OR YGL058W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85113143; PubMed=3881753;
RA Reynolds P., Weber S., Prakash L.;

```

R1 *RAD6 gene of Saccharomyces cerevisiae encodes a protein containing a
 R2 tract of 13 consecutive aspartates.*;
 R3 Proc. Natl. Acad. Sci. U.S.A. 82:168-172(1985).
 R4 [2]
 R5 SEQUENCE FROM N.A.
 R6 STRAIN S288c;
 R7 MEDLINE:97377993; PubMed:9234674;
 R8 Povermann M., de Montigny J., Fother S., Societ J.-L.;
 R9 "The characterization of two new clusters of duplicated genes
 R0 suggests a 'lego' organization of the yeast Saccharomyces cerevisiae
 R1 chromosomes.*;
 R2 Yeast 13:861-869(1997).
 R3 [3]
 R4 SOURCE OF 77 91, AND FUNCTION.
 R5 MEDLINE:87315384; PubMed:5306454;
 R6 Jenfisch S., McGrath J.P., Varshavsky A.;
 R7 "The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating
 R8 enzyme.*;
 R9 Nature 329:131-134(1987).
 R0 [4]
 R1 MUTAGENESIS OF CYS-68.
 R2 MEDLINE:90207263; PubMed:2157209;
 R3 Sunq P., Prakash S., Prakash S.;
 R4 "Mutation of cysteine-88 in the Saccharomyces cerevisiae RAD6 protein
 R5 abolishes its ubiquitin-conjugating activity and its various
 R6 biological functions.*;
 R7 Proc. Natl. Acad. Sci. U.S.A. 87:2695-2699(1990).
 R8 [5]
 R9 REQUIREMENT FOR E3.
 R0 MEDLINE:91293094; PubMed:2065660;
 R1 Sung P., Herleth E., Dickart C., Prakash S., Prakash S.;
 R2 "Yeast RAD6 encodes ubiquitin conjugating enzyme mediates protein
 R3 degradation dependent on the N end-recognizing E3 enzyme.*;
 R4 EMBO J. 10:2187-2193(1991).
 R5 [6]
 R6 X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 R7 MEDLINE:98165808; PubMed:9437453;
 R8 Wortylake D.K., Prakash S., Prakash S., Hill C.P.;
 R9 "Crystal structure of the Saccharomyces cerevisiae ubiquitin-
 R0 conjugating enzyme Rad6 at 2.6-A resolution.*;
 R1 J. Biol. Chem. 273:6271-6276(1998).
 R2 [7]
 R3 FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 R4 OTHER PROTEINS. UBC2 IS ACTIVE ON HISTONES. IT IS REQUIRED FOR
 R5 POSTREPLICATION REPAIR OF UV-DAMAGED DNA AND SPOROULATION. UBC2
 R6 MEDIATES E3-DEPENDENT UBC ACTIVITY.
 R7 CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
 R8 diphosphate + protein N-ubiquityllysine.
 R9 PATHWAY: Ubiquitin conjugation; second step.
 R0 SUBUNIT: Interacts with RAD6.
 R1 SUBCELLULAR LOCATION: Nuclear.
 R2 DOMAIN: THE ACIDIC-TAIL DOMAIN OF RAD6 IS IMPORTANT FOR
 R3 POLYUBIQUITINATION OF HISTONES, AS WELL AS FOR SPOROULATION.
 R4 MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 R5 UBIQUITIN-THIOLESTER FORMATION.
 R6 SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 R7
 R8 This SWISS-PROT entry is copyright. It is produced through a collaboration
 R9 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 R0 the European Bioinformatics Institute. There are no restrictions on its
 R1 use by non-profit institutions as long as its content is in no way
 R2 modified and this statement is not removed. Usage by and for commercial
 R3 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 R4 or send an email to license@sib-sib.ch).
 R5
 R6 EMBL: K02962; AAA34952.1;
 R7 EMBL: Z72580; CAA9761.1;
 R8 PIR: A21906; A21906.
 R9 PDB: 1AY7; 26-AUG-98.
 R0 SGD: S0003026; RAD6.
 R1 InterPro: IPR000608; UBC conjugat.
 R2 Pfam: PF00179; UBC.con; 1.
 R3 Prodom: PD000463; UBC conjugat; 1.
 R4 SMART: SM00212; UBCc; 1.

DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE; PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ubl conjugation pathway; Ligase; DNA repair; Sporulation;
 KW Nuclear protein; Multigene family; 3D-structure.
 FI DOMAIN 150 172 ASP/GLU-RICH (ACIDIC TAIL).
 FI MOD_RFS 1 1 BLOCKED.
 FT BINDING 88 88 UBIQUITIN.
 FT MUTAGEN 88 88 C->A,V; LOSS OF ACTIVITY.
 SQ SEQUENCE 172 AA; 19705 MW; 5F568DC28ABBD60F CRC64;
 Query Match 20.58; Score 310; DB 1; Length 172;
 Best Local Similarity 30.08; Pred. No. 4.9e-18;
 Matches 66; Conservative 46; Mismatches 58; Indels 50; Gaps 4;
 QY 7 ISSOKALMLELKSGLCEPVEGFRTILVDESGLYNEVAJGLPNTLYEGGYFAHAKFKFI 66
 DB 3 IPARRLLMRDFKMKEDAPGVSPSPDNVYV-NNAMTIGPADTPTEDATFRLLEFDE 61
 QY 67 DYPSPPTFRFLTKMMHPNIYENGDCVISTLHPVDDPQSGFLASFERNITQNVRTILLS 126
 DB 62 EYPNKPPIVKEFSEMFHPNVYANGEICLDL-----QNRKPTDYDVASILIS 108
 QY 127 VISLLNEPNTFSPANVDASVYFRKWRKSKOKKEYAEIIRKQVSATRAEAKGKGVKPTT 186
 DB 109 TSLFNDPNPASPANVEAATLF-----KDKSKQYVVRKRVETVRKSWEDM----- 153
 QY 187 LAEYCIKTKVPSNDNSLLYDDLYDDDDIDDEDEEEEDAD 226
 DB 154 -----DDMDDDDDDDDDDDDDDDDDDD 172

RESULT 14
 UBC2_CANAL STANDARD; FRT: 179 AA.
 AC 07420i;
 DI 15-JUL-1999 (Rel. 38, Created)
 DI 15-JUL-1999 (Rel. 38, Last sequence update)
 DI 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-20 kDa (EC 6.3.2.19)
 GN (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
 OS UBC2 OR RAD6.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3153;
 RA Leng P., Brown A.J.P.;
 RT "Molecular cloning and characterisation of CarA6 gene from Candida
 RL albicans.*;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3153A;
 RX MEDLINE=20177840; PubMed=10712766;
 RA Leng P., Sudbery P.E., Brown A.J.P.;
 RT "Radp represses yeast-hypha morphogenesis in the human fungal
 RI pathogen Candida albicans.*;
 RC Mol. Microbiol. 35:1264-1275(2000).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS. ALSO REGULATES YEAST-HYPHA MORPHOGENESIS.
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: AF046707; AAC24755.1;
 EMBL: AF118145; AAC45241.1;
 HSSP: P06104; JAY.
 InterPro: IPR000608; UBQ_conjugat.
 Pfam: PF00179; UBQ_conj.
 ProDom: P000461; UBQ_conjugat. 1.
 SMART: SM00212; UBQ.
 PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 PROSITE: PS50127; UBIQUITIN_CONJUGAT_2; 1.
 K1 Ub1 conjugation pathway; ligase; DNA repair; Nuclear protein;
 M. ligase family.
 FT DOMAIN 151 179 ACIDIC TAIL.
 FT BINDING 88 88 UBIQUITIN (HY SIMILARITY).
 SQ SEQUENCE 179 AA; 20576 MW; 24A081B5F8FAD316 CRC64;

Query Match 20.4%; Score 308; DB 1; Length 179;
 Best Local Similarity 30.1%; Pred. No. 7.5e-48;
 Matches 58; Conservative 48; Mismatches 54; Indels 56; Gaps 6;
 QY 7 TSSOKA:MLELKSQEPVGFRTILVDSLYNNEVA:EGLPNTLYEGYFKAHIKPPI 66
 DB 3 TVAREL:MRQFKRQKQFQSVASPIFD-NYKKNVILGPSDTPEDGTRILLSEDE 61
 QY 67 DYPSPPTFRELTKMHIPNIYENGVCISILHPVDDPQSGELPSRWNPNTONVITILLS 126
 DB 62 QYKPKPVKVFISEMHPNVAASGELCLDIL-----QNRKSPYDVSSITLS 108
 QY 127 VISLLNPNFSPANVDASVMFKRWSKGDKEYAEILRKQVSATKAEEKGVKVPPI 186
 DB 139 VSLNLDNPNSSPANVEANLYKHRSI-----YVKRVRETV----- 145
 QY 187 LAEYC:RTKVPNSDSD:ILLYDDLYDDIDDEDEEDADYDDSD 232
 DB 146 -----ENS-----WMD--DDEEEEDDEADDEDDDD 172

RESULT 15
 UBC2_HUMAN
 ID UBC2_HUMAN STANDARD; PRT; 192 AA.
 AC P52478; Q45062;
 DT 01-OCT-1995 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-21.5 kDa (EC 6.3.2.19) (Ubiquitin-
 protein ligase) (Ubiquitin carrier protein).
 GN UBC2 OR C35B1.1
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Poloderinae; Caenorhabditis.
 CC NCB_TaxID:6239;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2.
 RX MEDLINE:96027757; PubMed-7545294;
 RA Jorgensen D.S., Jones D., Candido E.P.M.;
 RT "Caenorhabditis elegans UBC-1, a ubiquitin-conjugating enzyme
 RT homologous to yeast RAD6/UBC2, contains a novel carboxy-terminal
 RT extension that is conserved in nematodes."
 RL DNA Cell Biol. 14:883-891(1995).
 RN 12
 RP SEQUENCE FROM N.A.
 RC S. RAIN-Bristol N2.
 RA Miller N., Stellys L., Bradshaw H., Keppeler D.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN 13
 RP REVISTONS.
 RA Waterston R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

CC CC FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC CC PROTEINS.
 CC CC CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
 CC CC diphosphate + protein N-ubiquityllysine.
 CC CC PATHWAY: Ubiquitin conjugation; second step.
 CC CC MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC CC UBIQUITIN-THIOLESTER FORMATION.
 CC CC SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC CC STRONGEST TO YEAST UBC2.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC CC or send an email to license@isb-sib.ch).

EMBL: U08139; AAA83388.1;
 EMBL: AF045638; AAC02561.2;
 HSSP: P25865; ZAAK.
 WormPep: C35B1.1; CE27822.
 InterPro: IPR000608; UBQ_conjugat.
 Pfam: PF00179; UBQ_conj. 1.
 ProDom: P000461; UBQ_conjugat. 1.
 SMART: SM00212; UBQ.
 PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 PROSITE: PS50127; UBIQUITIN_CONJUGAT_2; 1.
 K1 Ub1 conjugation pathway; Ligase; Multigene family.
 FT BINDING 88 88 UBIQUITIN (HY SIMILARITY).
 SQ SEQUENCE 192 AA; 21513 MW; 7CF268F8F956F33D CRC64;

Query Match 20.0%; Score 301.5; DB 1; Length 192;
 Best Local Similarity 31.4%; Pred. No. 2.7e-17;
 Matches 72; Conservative 42; Mismatches 72; Indels 43; Gaps 5;
 QY 7 TSSOKA:MLELKSQEPVGFRTILVDSLYNNEVA:EGLPNTLYEGYFKAHIKPPI 66
 DB 3 TPRSRLMRQFKLQEDPPAGVS-GAPTEINLTWEALIFGQEIFEDSTFKLSLETE 61
 QY 67 DYPSPPTFRELTKMHIPNIYENGVCISILHPVDDPQSGELPSRWNPNTONVITILLS 126
 DB 62 SYNPKPPTVKFISKMEHPNVAASGELCLDIL-----QNRKSPYDVAAITLS 108
 QY 127 VISLLNPNFSPANVDASVMFKRWSKGDKEYAEILRKQVSATKAEEKGVKVPPI 186
 DB 139 IQSLDEPNPNSPANSLAQIYQENR-----REYKRVQOIVE----- 145
 QY 187 LAEYC:RTKVPNSDSD:ILLYDDLYDDIDDEDEEDADYDDSD 234
 DB 147 -----QSWLNFGENGDVAKDDVEIEEIAAPGANDADDDDEKASG 189

RESULT 16
 UBC2_HUMAN
 ID UBC2_HUMAN STANDARD; PRT; 152 AA.
 AC P23567;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2 B (EC 6.3.2.19) (Ubiquitin-protein
 DE ligase B) (Ubiquitin carrier protein B) (HR6B) (HR6B) (F2-17 kDa).
 GN UBE2B OR RAD6B
 OS Homo sapiens (Human).
 OS Mus musculus (Mouse).
 OS Rattus norvegicus (Rat), and
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC NCB_TaxID:9606, 10090, 10116, 9986;
 RN 11
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC SPECIES-Human;

RX MEDLINE-90228340: PubMed-2158443;
 RA Schneider M., Beckers Korn C., Lottspeich F., Schweiger M.;
 RI The human ubiquitin carrier protein E2(MR-17,000) is homologous to
 RC the yeast DNA repair gene RAD6.*;
 RN EMBL J. 9:1431-1435(1990).
 RP SPECIES-HUMAN.
 RX MEDLINE-92020951: PubMed-1717990;
 RA Koken M.H.M., Reynolds P., Jaspers-Dekker I., Prakash L., Prakash S.,
 KA Bootsma D., Hoelmakers J.H.J.;
 RT "Structural and functional conservation of two human homologs of the
 RT yeast DNA repair gene RAD6.*;
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8865-8869(1991).
 RN [3]
 RP SPECIES-HUMAN.
 RX MEDLINE-92020951: PubMed-1717990;
 RA Poljimenko A., Blocher S.R.;
 RI "Exon-intron structure of mammalian HRG6 and HR68 genes.*;
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SPECIES-HUMAN.
 RX MEDLINE-92020951: PubMed-1717990;
 RA Wang S.S., Dumas F., Banville D.;
 RI "A rabbit reticulocyte ubiquitin carrier protein that supports
 RT ubiquitin-dependent proteolysis (E214k) is homologous to the yeast
 RT DNA repair gene RAD6.*;
 RL J. Biol. Chem. 267:6495-6501(1992).
 RN [5]
 RP SPECIES-HUMAN.
 RX MEDLINE-94324482: PubMed-8048511;
 RA Wang S.S., Banville D.;
 RI "14 kDa ubiquitin-conjugating enzyme: structure of the rat gene and
 RI regulation upon fasting and by insulin.*;
 RL Am. J. Physiol. 267:E39-E48(1994).
 RN [6]
 RP SPECIES-HUMAN.
 RX MEDLINE-94324482: PubMed-8048511;
 RA Roest H.P., van Klaveren J., de Wit J., van Gurp C.G., Koken M.H.M.,
 KA Vermeij M., van Rijen J.H., Vreuburg J.T.M., Haerends W.M.,
 KA Bootsma D., Grootenboer J.A., Hoelmakers J.H.J.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SPECIES-HUMAN.
 RX MEDLINE-94324482: PubMed-8048511;
 RA Varshavsky A., Grigoriyev S., Stewart A.E., Kwon Y.T., Arfin S.M.,
 RA Bradshaw R.A., Jenkins N.A., Copeland N.G.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC 1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. REQUIRED FOR POSTREPLICATION REPAIR OF UV-DAMAGED
 CC DNA.
 CC 1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC 1- PATHWAY: Ubiquitin conjugation; second step.
 CC 1- SUBUNIT: Interacts with RAD8.
 CC 1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC 1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC STRONGEST: TO YEAST UBC2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M74525; AAA35982.1;
 DR EMBL: X53251; CAA33739.1;
 DR EMBL: AF294392; AAG27628.1;

DR EMBL: AF294387; AAG27628.1; JOINED.
 DR EMBL: AF294388; AAG27628.1; JOINED.
 DR EMBL: AF294389; AAG27628.1; JOINED.
 DR EMBL: AF294390; AAG27628.1; JOINED.
 DR EMBL: AF294391; AAG27628.1; JOINED.
 DR EMBL: M62387; AAA31452.1;
 DR EMBL: M62388; AAA21087.1;
 DR EMBL: U04308; AAB60669.1;
 DR EMBL: U04303; AAB60669.1; JOINED.
 DR EMBL: U04304; AAB60669.1; JOINED.
 DR EMBL: U04305; AAB60669.1; JOINED.
 DR EMBL: U04306; AAB60669.1; JOINED.
 DR EMBL: U04307; AAB60669.1; JOINED.
 DR EMBL: X96859; CAA65602.1;
 DR EMBL: U57690; AAC52884.1;
 DR PIR: S12530; S12530.
 DR PIR: B41222; B41222.
 DR PIR: A42416; A42416.
 DR HSSP: P25865; ZAAK.
 DR Genew: HGNC:12473; UBE2B.
 DR MIM: 179095;
 DR MIM: 102944; UBE2b.
 DR InterPro: IPR006608; UBC_conjugat.
 DR Pfam: PF00179; UBC_conjugat; 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBC2; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ubl conjugation pathway; Ligase; DNA repair; Nuclear protein;
 KW Multigene family.
 FT BINDING 88 UBIQUITIN (HY SIMILARITY).
 SQ SEQUENCE 152 AA; 17312 MW; CDFEERTEU6840BE CRC64;
 Query Match 19.5%; Score 295; DB 1; Length 152;
 Best Local Similarity 37.6%; Pred. No. 6.6e-17;
 Matches 65; Conservative 31; Mismatches 53; Indels 24; Gaps 4;
 QY 7 TSSKALMLELKSLEQPEVGEFRITLVDSGLYNWEVAIFGLPNTLYEGGYFKAHKKFPI 66
 DB 3 TPARRLRMRQKRLQEDPEVGS-CAPEENNINQWNAVIFGPEPTDGTFLVTFESE 61
 QY 67 DYPYSPPTFFFLFKMHPNLYENGDCVLSILHPVDPDPOSGELPSEHNPNTQNVRTILLS 126
 DB 62 EYPNKPPTVRFELSKMHPNLYADSCIDIL-----QNRWSPDYDSSILIS 108
 QY 127 VTSLNENPTFSPANDVSWFKRWKROSKGKDEYAEIIRKQVSAIKAEKD 179
 DB 109 TQSLLDEPNPNSPANSOAAQIYCE-----NKRKY-----FKRVAIVHQSWND 151
 RESULT 17
 ID UBCA_HUMAN STANDARD: PRT; 152 AA.
 AC P49459;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2A (EC 6.3.2.19) (Ubiquitin-protein
 DE ligase A) (ubiquitin carrier protein A) (UBCA) (HHR6A).
 GN UBCA OR RAD6A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SPECIES FROM N.A.
 RX MEDLINE-92020951: PubMed-1717990;
 RA Koken M.H.M., Reynolds P., Jaspers-Dekker I., Prakash L., Prakash S.,
 KA Bootsma D., Hoelmakers J.H.J.;
 RI "Structural and functional conservation of two human homologs of the
 RI yeast DNA repair gene RAD6.*;
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8865-8869(1991).
 RN [2]

Query Match	18.5%	Score 261:	26 1:	Length 151:
Best local Similarity	35.0%	Prod. No. 8.6e-16:		
Matches	57:	Conservative	36:	Mismatches 50: Indels 20: Gaps
QY	7	TSOKALELKSQEEPVGEFRITLVDESOLYWEVAIFGLPNTLYEGGYFAHAKFPI	66	
DB	3	TAARRRI.MRDEKMTDPPAGVSASHPVD-NVMTNVALIGPADTFEDCHFRIVMHFE	61	
QY	67	DYPSPTPTFRFTKMHNPNIYENGVCISILHPVDPOSGGELPSERNWPNQVRIILLS	126	
DB	62	QYNKPSPSYKVFISEFNPYATGELCIDL-----QNRMSPTYDVAALVLS	108	
QY	127	VISILNEPNTSPANVDIASYMKRKRDSKGGKKEYAFIRKQV	169	
DB	109	IQSLLNDPNTGSPANVFSNLY-----KDRKEYKRVRETV	145	
RESULT 20				
QY	UHC2_ARATH	STANDARD;	PRT;	152 AA.
DB	P42745:			
DT	01-NOV-1995	(Rel. 32, Created)		
DT	01-NOV-1995	(Rel. 32, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Ubiquitin-conjugating enzyme E2-17 kDa 2 (EC 6.3.2.19) (Ubiquitin-			
DE	protein ligase 2) (Ubiquitin carrier protein 2).			
GN	UBC2 CR AT2602760 OR T20F6.10.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid 1; Brassicales; Brassicaceae; Arabidopsi.			
OX	NCBI_TaxID:3702;			
RT	[1]			
RT	SEQUENCE FROM N.A.			
RC	STRAIN-cv. Columbia; TISSUE-Green leaf;			
RC	MDLINC=94207190; PubMed=8155884;			
RX	Sullivan M.L., Carpenter I.B., Vierstra R.D.:			
RT	"Homologues of wheat ubiquitin-conjugating enzymes-- laURC1 and			
RT	laUBC4 are encoded by small multigene families in Arabidopsis			
RT	thaliana."			
RT	Plant Mol. Biol. 24:651-661(1994).			
RN	[2]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-cv. Columbia;			
RP	MDLINC=98041921; PubMed=9371881;			
RC	Wahr P., Stary S., Luschig C., Bachmair A.:			
RT	"Arabidopsis thaliana RAD6 homolog ARUBC2 complements UV sensitivity,			
RT	but not N-end rule degradation deficiency, of Saccharomyces			
RT	cerevisiae rad6 mutants."			
RI	Curr. Genet. 32:309-314(1997).			
RL	[3]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-cv. Columbia;			
RP	MDLINC=20083487; PubMed=10617197;			
RC	Lie X., Kaul S., Rounsley S.D., Shea J.P., Benito M.-J., Town C.D.,			
RA	Fujii C.Y., Mason I.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,			
RA	Buell C.R., Ketchum K.A., Lee J.-J., Rinning C.M., Koo H.L.,			
RA	Moffat L.S., Cronin L.A., Shen M., Pail G., Van Aken S., Umayam L.,			
RA	Tallon K.J., Gill J.E., Adams M.B., Carrera A.J., Creasy T.H.,			
RA	Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,			
RA	Nierman W.C., White O., Kisen J.A., Salzberg S.L., Fraser C.M.,			
RA	Venter J.C.:			
RT	"Sequence and analysis of chromosome 2 of the plant Arabidopsis			
RT	thaliana."			
RT	Nature 402:761-768(1999).			
CC	-1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO			
CC	OTHER PROTEINS. REQUIRED FOR POSTREPLICATION REPAIR OF UV-DAMAGED			
CC	DNA (BY SIMILARITY).			
CC	-1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +			
CC	diphosphate + protein N-ubiquityllysine.			
CC	-1- PATHWAY: Ubiquitin conjugation; second step.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED. LOWER			
CC	LEVELS FOUND IN LEAVES.			

RA Berkova D., Botchan M.R., Brock J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Catter A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Cavenport L.B., Davies P.,
RA De Pablo S., Delcher A., Deng Z., Dew A.D., Dew I., Dietz S.M.,
RA Dodson K., Dop L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsin C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell C.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Hejman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Hwang S.H., Ibeqam C.,
RA Jallali M., Kaulsh F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li X., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh L.C., McLeod M.P., Mohrson D.,
RA Merkley G., Milshina N.V., Mobarry C., Morris J., Mostrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson C.K., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzo G.M., Pittman C.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shih H.C., Sider-Kimms I., Simpson M., Skupski M.P., Smith H.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zheng F.N., Zheng W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of *Drosophila melanogaster*."
RA Science 287:2185-2195(2000).
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC DNA.
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine + AMP ->
CC diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC -!- STRONGEST: TO YEAST RAD6.

RA FMBB: ME3792; AAA28109.1;
RA EMBL: ME4435; AAA28109.1;
RA EMBL: ME3791; AAA28109.1; JOINED.
RA EMBL: AF003604; AAF52079.1;
RA PIR: A39392; A39392.
RA F589; PZ5865; 2AAB.
RA Flybase: FBgn004436; Ubq26.
RA InterPro: IPR000608; Ubq-conjugat.
RA Pfam: PF00179; Ubq-con; 1.
RA Prodom: PB003461; Ubq-conjugat; 1.
RA SMART: SM00212; UBQ; 1.
RA PROSITE: PS00183; UBIQUITIN-CONJUGAT-1; 1.
RA PROSITE: PS00127; UBIQUITIN-CONJUGAT-2; 1.
RA Ub1 conjugation pathway; Ligate; DNA repair; Nuclear protein;
RA Multigene family.
RA BONDING 88 88 UBIQUITIN (BY SIMILARITY).
RA CONFID 95 99 T -> R (IN REF. 1).
RA SEQUENCE 151 AA; 17152 MW; CC4B3592E4A9220 CRC64;
Query Match 18.3%; Score 277; DR 1; Length 151;
Best Local Similarity 36.2%; Pred. No. 1.8e-15;
Matches 58; Conservative 34; Mismatches 48; Indels 20; Gaps 3;

QY 7 TSSQAM:ELKSLQRPVSPCHITINDSLYKNEVAIFCLPNTLYEGYKAHKPKPI 66

DB 3 IPARRLMRFKKLQDDPTGVSAPTID-NNIMLWNAVTFGRDTPFDGFKLTIFTE 61
QY 67 DYKSPPTFRFLTKMHPNIYENGWQVIGISILHPEVDPOSGELPSERNWNTONVRIILS 126
DB 62 EYNPKPPTVRFVSKVHPNVAAGGICLDIL-----QNRWSPYDVSAIILS 108
QY 127 VISLNEPNTFFSPANDVASVMPKRWKSKDKKPYAHILR 166
DB 109 TQSLSDPNPNSPANSTAAGLYKENR-----REYKRVK 142

RESULT 23
UBCL_ARATH
ID UBCL_ARATH STANDARD: PRT: 152 AA.
AC P25865;
DI 01-MAY-1992 (Rel. 22, Created)
DI 01-MAY-1992 (Rel. 22, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 1 (EC 6.3.2.19) (Ubiquitin-
DE protein ligase 1) (Ubiquitin carrier protein 1).
GN UBCL OR AT1G14400 OR F14L17_35 OR F14L17.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=92084684; PubMed=1660887;
RI Sullivan M.L., Vierstra R.D.;
RI "Cloning of a 16-kDa ubiquitin carrier protein from wheat and
RI Arabidopsis thaliana. Identification of functional domains by in
RI vitro mutagenesis.";
RI J. Biol. Chem. 266:23878-23885(1991).
RN 12;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=94207190; PubMed=8155884;
RI Sullivan M.L., Carpenter T.B., Vierstra R.D.;
RI "Homologues of wheat ubiquitin-conjugating enzymes -- UbHCL and
RI TaUBC4 are encoded by small multigene families in Arabidopsis
RI thaliana.";
RI Plant Mol. Biol. 24:651-661(1994).
RN 13;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RI Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RI White O., Alonso J., Aitafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RI Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RI Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RI Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RI Gill J.E., Goldsmith A.D., Haas R., Hansen N.P., Hughes B., Huizar L.,
RI Hunter J.I., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RI Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RI Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RI Lin X., Liu S.X., Liu Z.A., Luros J.S., Maili R., Marzilli A.,
RI Milišcher J., Miranda M., Nguyen M., Nicrman W.C., Osborne B.I.,
RI Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RI Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RI Sun H., Taiton L.J., Tambunga G., Torkum M.J., Town C.D.,
RI Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RI Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RI "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RI thaliana.";
RI Nature 408:816-820(2000).
RN 14;
RP SEQUENCE OF 1-98 FROM N.A.
RC STRAIN=cv. Columbia;
RI Keynal M., Grollet F., Landic M., Meyer Y., Cooke R., Nelson M.;
RI Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.

[5]
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RP MEDLINE:92340563; PubMed-1321826;
 RA Cook W.J., Jeffrey L.C., Sullivan M.L., Vierstra R.D.;
 RI "Three-dimensional structure of a ubiquitin-conjugating enzyme (E2).";
 RL J. Biol. Chem. 267:15116-15121(1992).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. REQUIRED FOR POSTREPLICATION REPAIR OF UV-DAMAGED
 CC DNA (HY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin -> protein lysine - AMP +
 CC diphosphate - protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M62721; AAA32303.1; -;
 DR EMBL: L49351; AAA32897.1; -;
 DR EMBL: AC012188; AAF43940.1; -;
 DR EMBL: Z27262; CA881773.1; -;
 DR PDB: 2AAK; 18-MAR-98.
 DR InterPro: IPR000608; Ubq_conjugat.
 DR Pfam: PF00179; Ubq_conj; 1.
 DR ProDom: PD000461; Ubq_conjugat; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ubl conjugation pathway; Ligase; Multigene family; 3D-structure.
 FT BINDING 88 88 UBIQUITIN (BY SIMILARITY).
 FT TURN 5 7
 FT HELIX 8 16
 FT TURN 17 18
 FT TURN 22 23
 FT STRAND 24 29
 FT TURN 30 31
 FT STRAND 32 42
 FT TURN 44 45
 FT TURN 47 48
 FT STRAND 49 49
 FT TURN 50 50
 FT STRAND 52 58
 FT TURN 61 65
 FT STRAND 69 72
 FT TURN 79 80
 FT STRAND 81 81
 FT TURN 83 84
 FT STRAND 87 87
 FT HELIX 90 93
 FT TURN 94 94
 FT TURN 98 99
 FT TURN 102 114
 FT TURN 118 119
 FT HELIX 124 132
 FT HELIX 134 147
 FT TURN 148 149
 SQ SEQUENCE 152 AA: 17281 MW: 8961911930C99045 CRC64;

Query Match 18.3%; Score 276; DB 1; Length 152;
 Best Local Similarity 36.8%; Pred. No. 2, 2e-15;
 Matches 60; Conservative 32; Mismatches 51; Indels 20; Gaps 4;

QY 7 TSSKALMLKSLQEEVPEFRITIVDESILYNNMEVAIFGLPNTLYEGGYFKAHLKFT 66
 DB 3 TPARKRLMRDKRLQDDPPA31S-GAPQDNNIMLNNAVFGLPDDTPWDGCTFKLSLPSE 61

QY 67 DYPYSPPTFRFLTKMHPNIYENGVCISILHPPVDVDFQSGELPSERWNPNTQVTRILLS 126
 DB 62 DPNKPPPTREVSRMHPNIYADGSGICIDIL-----QNWSPHYDVAAILIS 108
 QY 127 VTSLLNEPNI*SPANVDASVMFKWHDGKDKVKAHLIRKQV 169
 DB 109 IOSLLCDPNRSPANSENA---RMYSESK---REYNRRVRDVV 145
 RESULT 24
 URC2.WHEAT
 ID UBC2.WHEAT STANDARD; PRT; 152 AA.
 AC P25866;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
 DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
 GN UBC1
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-88.
 RC STRAIN=cv. Augusta;
 RX MEDLINE:92084684; PubMed-1660887;
 RA Sullivan M.L., Vierstra R.D.;
 RT *Cloning of a 16-kDa ubiquitin carrier protein from wheat and
 RT Arabidopsis thaliana. Identification of functional domains by in
 RT viro mutagenesis.;
 RL J. Biol. Chem. 266:23878-23885(1991).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS.
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M62720; AAA34310.1; -;
 DR HSSP: P25865; 2AAK.
 DR InterPro: IPR000608; Ubq_conjugat.
 DR Pfam: PF00179; Ubq_conj; 1.
 DR ProDom: PD000461; Ubq_conjugat; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ubl conjugation pathway; Ligase; Multigene family.
 FT BINDING 88 88 UBIQUITIN.
 FT MUTAGEN C->S: LOSS OF ACTIVITY.
 SQ SEQUENCE 152 AA: 17301 MW: F13AA5439EB4C8AB CRC64;

Query Match 18.3%; Score 276; DB 1; Length 152;
 Best Local Similarity 37.1%; Pred. No. 2, 2e-15;
 Matches 63; Conservative 29; Mismatches 44; Indels 34; Gaps 5;

QY 7 TSSKALMLKSLQEEVPEFRITIVDESILYNNMEVAIFGLPNTLYEGGYFKAHLKFT 59
 DB 3 TPARKRLMRDKRLQDDPPA31S-GAPQDNNIMLNNAVFGLPDDTPWDGCTFKLSLPSE 54
 QY 60 AILKEPTIDYSPPTFRFLTKMHPNIYENGVCISILHPPVDVDFQSGELPSERWNPNTQ 119
 DB 109 IOSLLCDPNRSPANSENA---RMYSESK---REYNRRVRDVV 145

DB 55 LT:QFTEYVKNPPTVREVSMEHFN:YADGSLICLIL-----ONOWSDIYD 101

QY 120 VRTILLSVLSLNEPTSPANVDASVWFRKRGSKGKDYAEIRKQV 169

DB 102 VAAITISIGSLICLIDPNPNSPANSFARMY-----SENK-REYNKRVREV 145

RESULT 25

UBCD_SCHPO STANDARD; PRT: 176 AA.

ID UBC3_ARATH

AC 000103:

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ubiquitin-conjugating enzyme E2-20 kDa (EC 6.3.2.19)

DE (Ubiquitin protein ligase) (Ubiquitin carrier protein).

GN UBCP4 OR SPEC1259, 19C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomyces.

OX NCBI_TaxID=4896;

RN SEQUENCE FROM N.A.

RP STRAIN=972:

RC SRA:IN-972:

RX MEDLINE=97299689; PubMed=9154838;

RA Osaoka F., Sano T., Sano T., Yamao F.;

RA "A ubiquitin-conjugating enzyme in fission yeast that is essential for the onset of anaphase in mitosis.";

RI Mol. Cell. Biol. 17:3388-3397(1997).

RN 121

RP SEQUENCE FROM N.A.

RC STRAIN=972:

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgoutos J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor K., Cronin A., Davis P., Feltwell I., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby I., Howarth S., Huckle E.J., Hunt S., Jagals K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch F.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skellon J., Simmonds M., Squares R., Stevens K., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Woodward J., Vackaert G., Aert R., Robben J., Grymoprez B.,

RA Welljones L., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritze C., Holzer F., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Leirach H., Weinhardt R., Pohl T.M.,

RA Eder P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu F., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

RA Page R.K., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.F., Moreno S., Armstrong J., Forsberg S.L.,

RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Supakowski G.V., Ussery D., Barrell B.G., Nurse P.;

RI "The genome sequence of Schizosaccharomyces pombe.";

RN Nature 415:871-880(2002).

CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS.

CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP + diphosphate + protein N-ubiquityllysine.

CC -1- PATHWAY: Ubiquitin conjugation; second step.

CC -1- MISCELLANEOUS: A CYSINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOLESTER FORMATION.

CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.

CC STRONGEST, TO YEAST UBC11.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR HMBL: D85545; HAA20375.1; -.

DR EMBL: AL034564; CAB38416.1; -.

DR HSSP: Q95044; 2F2C.

DR InterPro: IPR000608; UBO_conjugat.

DR Pfam: PF00179; UO_con: 1.

DR ProDom: PD000461; UBO_conjugat; 1.

DR SMART: SM00212; UBCC; 1.

DR PROSITE: PS00183; UBIQUITIN_CONJUGAT.1; 1.

DR PROSITE: PS01027; UBIQUITIN_CONJUGAT.2; 1.

KW Ubl conjugation pathway; Ligase; Mitosis; Cell cycle; Cell division;

KW Multigene family. 113

FT BINDING 113 UBIQUITIN (BY SIMILARITY).

SQ SEQUENCE 176 AA; 19664 MW; 6C80769C37AD4950 CRC64;

Query Match 17.7%; Score 267.5; DH 1; Length 176;

Best Local Similarity 39.9%; Pred. No. 1.2e-14;

Matches 57; Conservative 20; Mismatches 53; Indels 13; Gaps 1;

QY 8 SSQAUMLEKLSIQEPEVGFRIILVDESILYNEVAIFGLPNTIYEGCYFKAHKEP:D 67

DB 28 SVTKRLKSELMSLMSNTPTCSAIPDSDSLHLHWAQITITGSDTYEGIKFKISMSFPAN 87

QY 68 YPYSPPTFRFLTMHHPNIYENGVCISILHPVDPQSGELPSEKWNPTONVITLLSV 127

DB 86 YPYSPPTFRITFTSPMHPNVDMSGNTCLDIL-----KOKWSAVYVQVITLLSL 134

QY 128 ISHLNEPNTFSPANVDASVWFRK 150

DB 135 QSLIGEPNNSPLNQAQAEELSK 157

RESULT 26

ID UBC3_ARATH STANDARD; PRT: 150 AA.

AC P42746:

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ubiquitin-conjugating enzyme E2-17 kDa 3 (EC 6.3.2.19) (Ubiquitin-protein ligase 3) (Ubiquitin carrier protein 3).

GN UBC3 OR A15G62540 OR K19B1.15.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OX eurosids II; Brassicales; Brassicaceae; Arabidopsids.

RN NCBI_TaxID=1702;

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia; TISSUE=Green leaf;

RX MEDLINE=94207190; PubMed=8155884;

RA Sullivan M.L., Carpenter I.B., Vierstra R.D.;

RT "Homologues of wheat ubiquitin-conjugating enzymes -- TaUBC1 and TaUBC4 are encoded by small multigene families in Arabidopsis thaliana.";

RT Plant Mol. Biol. 24:651-661(1994).

RN 121

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=99087489; PubMed=9872454;

RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N., Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII. Sequence features of the regions of 1,013,767 bp covered by sixteen physically assigned pl and TAC clones.";

RL DNA Res. 5:297-308(1998).

CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS. REQUIRED FOR POSTREPLICATION REPAIR OF UV-DAMAGED DNA (BY SIMILARITY).

DR PIR: S32673; S32673.
 DR HSP: P1573; IQCG.
 DR InterPro: IPR006068; UBC_conjugat.
 DR Pfam: PF00179; UQ_con: 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM0212; UBCG; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_2; 1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_1; 1.
 KW Ubl conjugation pathway; Ligase; Multigene family.
 FT BINDING 85 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 148 AA: 15551 MW: 74A0991115BA50EA CRC64;
 Query Match 17.3%; Score 261; DB 1; Length 148;
 Best Local Similarity 36.2%; Pred. No. 3.3e-14;
 Matches 51; Conservative 25; Mismatches 51; Indels 14; Gaps 2;
 QY 9 SOKALMELKSLQPEVEGRITVDSELYNNEVAIFG:PNLYEGGYEKAH:KEPIDY 68
 ID 2 ASKRLLEADQKQVPSNCAGVAE:DMHRCATIMPPSPYAGGVFVSHPPDY 60
 QY 69 PYSPTTFKTKMHPNTYVNDGVCISILHPVDEPQSGELSPERNPTONVTRILLSVI 128
 DB 61 PIKEKVSRTKYVHPNLSNGSICLDIL-----KEQSPALITKVLISIC 107
 QY 123 SLI NEPTFSPANVDSYWER 149
 DB 138 SLITDPAQCPVPEFAIMYK 128
 RESULT 30
 UBC3_SCHPO STANDARD: PRT; 157 AA.
 AC P40944;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-18 kDa (EC 6.3.2.19)
 DE (Ubiquitin-conjugating enzyme HUS5) (Ubiquitin carrier protein HUS5).
 GN HUS5 OR SPAC30D11.13.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CX NCBI_LtaxID 4890;
 RN [1].
 RP SEQUENCE FROM N.A.
 RX MEDLINE:95286695; PubMed-7769995;
 RA Al-Khodairy F., Enoch T., Hagman J.M., Carr A.M.;
 RI "The Schizosaccharomyces pombe hus5 gene encodes a ubiquitin
 J. Cell Sci. 108:475-486(1995).
 RN [2].
 RP SEQUENCE FROM N.A.
 RX STRAIN-972;
 RA Wood V., Gwilliam K., Rajandream M.A., Lyne M., Gyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell I., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby I., Howarth S., Hickie E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Mouton S., Murgall K., Murphy L., Niblett D., Odeh C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Butler S., Saunders D., Seeger K., Sharp S.,
 RA Skeator J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wooldraft G., Aert K., Robber J., Grynoprav B.,
 RA Welljens E., Vansurels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz G., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambuti R., Pernelle B.,
 RA Cottrill E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huist S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 RT The genome sequence of Schizosaccharomyces pombe.*;
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. REQUIRED FOR EFFICIENT RECOVERY FROM DNA DAMAGE OR
 CC S-PHASE ARREST AND NORMAL MITOSIS. THIS MAY BE AS PART OF A
 CC CHECKPOINT INDEPENDENT RECOVERY PROCESS.
 CC -!- CATALYTIC ACTIVITY: AIP + ubiquitin + protein lysine -> AMP +
 CC diisophosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC STRONGEST, TO YEAST UBC9.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcement/)
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: X81846; CAAS7438.1;
 DR EMBL: Z67961; CAAG1899.1;
 DR HSP: P50550; UUA9.
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF00179; UQ_con: 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM0212; UBCG; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ubl conjugation pathway; Ligase; DNA repair; Nuclear protein;
 KW Multigene family.
 FT BINDING 93 93 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 157 AA: 17989 MW: 4BF3A6389203D27F CRC64;
 Query Match 17.2%; Score 260.5; DB 1; Length 157;
 Best Local Similarity 40.6%; Pred. No. 3.9e-14;
 Matches 54; Conservative 20; Mismatches 40; Indels 19; Gaps 3;
 QY 37 DLYNNEVA:PLPNTLYEGGYEKAHKEPIDYSPPTTFKTKMHPNTYVNDGVCISI 96
 DB 37 DLMNKKVGPCKPKTSWEGGLYKLMAPRYPTPRPKCHETPTPLHPNVPSCVCLSI 96
 QY 97 LHPVPDQSGELSPERNPTONVTRILLSVILLISVLLNPNTPSPANVDASVMPKRWKSG 156
 DB 97 LN-----FEEGKPKATIIKQIILGIQDILDPNTASPAQTEATYINPK----- 138
 QY 157 KDK:RYAETIRKQ 168
 DB 139 KQVEYKRVRAQ 151
 RESULT 31
 UBCG_HUMAN STANDARD: PRT; 179 AA.
 AC 000752;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2 C (EC 6.3.2.19) (Ubiquitin-protein
 DE ligase C) (Ubiquitin carrier protein C) (UBCH10).
 GN UB22C OR UBCH10.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CX NCBI_taxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND MUTAGENESIS OF CYS-114.
 RX MEDLINE:97225958; PubMed:912200;
 RA Townsley F.M., Arslanarkhov A., Beck S., Herskho A., Ruderman J.V.;
 RT Dominant-negative cyclin-selective ubiquitin carrier protein E2-
 C/Ubch10 blocks cells in metaphase;
 RL Proc. Natl. Acad. Sci. U.S.A. 94:2362-2367(1997).
 RP SEQUENCE FROM N.A.
 RX MEDLINE:21638749; PubMed:11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beadling L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
 RA Buck D.J., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.F., Corby N.K.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lechavalier M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie I.J., McEay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce I.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.L., Scott C.E., Schra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracy A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RI "The DNA sequence and comparative analysis of human chromosome 20";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS. REQUIRED FOR THE DESTRUCTION OF MITOTIC CYCLINS.
 CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -1- PATHWAY: Ubiquitin conjugation; second step.
 CC -1- MISCELLANEOUS: A CYS-114 RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC STRONGEST, TO YEAST UBC1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U73379; AAB53362.1;
 DR EMBL: AL050348; CAB66118.1;
 DR HSSP: Q95044; 2E2C.
 DR Genew: H0NC; 15937; UBE2C.
 DR MIM: 605574;
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBC; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT.1; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT.2; 1.
 DR UBI conjugation pathway; Ligase; Multigene family; Mitosis;
 KW Cell cycle; Cell division.
 FT BINDING 114 114 UBIQUITIN.
 FT VARIANT 25 25 G -> D.
 FT MUTAGEN 114 114 C->S; INHIBITION OF CYCLIN B DEGRADATION.
 FT
 SQ SEQUENCE 179 AA; 19652 MW; 0B6F58A1P0665D9A CRC64;
 Query Match 16.8%; Score 254.5; DB 1; Length 179;
 Best Local Similarity 37.0%; Pred. No. 1.4e-13;
 Matches 60; Conservative 25; Mismatches 58; Indels 19; Gaps 4;
 QY 11 KALMLEKLSLOEPPVPGFRITLVDSKLYNMEVAIFGLPNTLYEGGYFAHKEPPIIDY 69
 DB 33 KRLQQLMTLMMSGDKG--ISAFPESDNLKFWGTVIIGAAGTIVEDLRYKLSLEPSPGYP 90
 QY 70 YSPPTFRFLTKMHPNIYENGDVCIISILHPPVDVDSQSGELPSERNPNTQNYRTILLISVIS 129
 DB 91 YNAPIVAFILTPCHPNDVTCQGNICLDII-----KKWSALYDVRILLCSIOS 137
 QY 130 LLNEPNTFSPANDVASVFKWRUSKSKDKEYAEIIRKQVSA 171
 DB 138 LLGEPNIDSLPNTAAE--WKNPFAKKYLOETYSKQVTS 176
 RESULT 32
 UBC4_LYCES
 ID UBC4_LYCES STANDARD; PRT: 148 AA.
 AC P35135;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19) (Ubiquitin-
 DE protein ligase) (Ubiquitin carrier protein).
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridaceae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUP-Root;
 RA Bird D.M., Wilson M.A.;
 RL Submitted (XXA-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED
 CC AND ABNORMAL PROTEINS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -1- PATHWAY: Ubiquitin conjugation; second step.
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L23762; AAA34125.1;
 DR HSSP: P15731; IQCQ.
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBC; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT.1; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT.2; 1.
 KW Ubi conjugation pathway; Ligase; Multigene family
 FT BINDING 85 85 UBIQUITIN (BY SIMILARITY).
 FT RINDING 148 AA; 16522 MW; 4247FF721B98059F CRC64;
 SQ SEQUENCE 148 AA; 16522 MW; 16.7%; Score 252; DB 1; Length 148;
 Query Match 16.7%; Score 252; DB 1; Length 148;
 Best Local Similarity 35.5%; Pred. No. 1.7e-13;
 Matches 50; Conservative 26; Mismatches 51; Indels 14; Gaps 2;
 QY 9 SOKALMCELKSLQEEPFGFRITLVDSKLYNMEVAIFGLPNTLYEGGYFAHKEPPIIDY 68

```

DB 2 ASKRILKELKQKQDPPTSCSAGPVAE-DMFHQWATFMGPTDSDPYAGGVFLVTHRPDY 60
QY 59 PYSPTPTRELKMMHPNINYNJWCISILHPPVDDQSGELPSEKWNPTONVTRTILSVI 128
DB 61 PFAPPKVAERIKVFPNINSNCSCIDHIL-----KEQWSPALITSKVLISLC 107
QY 129 SILNENITSPANVDSWFER 149
DB 108 SLTDFNPDPLVPEIAIMYK 128

RESULT 33
LOC9_AWAT8
ID LOC9_AWAT8 STANDARD; PRT: 148 AA.
AC P35132: Q42014.
DT 01-FEB-1994 (Rel. 28, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 9 (EC 6.3.2.19) (Ubiquitin-
protein ligase 9) (Ubiquitin carrier protein 9) (UBCAT4B).
GN UBC9 OR UBC4B OR ATG527960 OR T1308.70.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids 11; Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=94359481; PubMed=8078482;
RA Genschik P., Burr A., Fleck J.
RA Giron P.-A., Carpenter T.B., van Nocker S., Sullivan M.L.,
RA Vierstra R.D.,
RT "Homologs of the essential ubiquitin conjugating enzymes UBC1, 4, and
RT 5 in yeast are encoded by a multigene family in Arabidopsis
RT thaliana."
RL Plant J. 3:545-552(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=94359481; PubMed=8078482;
RA Genschik P., Burr A., Fleck J.
RA "Differential expression of several E2-type ubiquitin carrier protein
RA genes at different developmental stages in Arabidopsis thaliana and
RA Nicotiana sylvestris."
RL Mol. Gen. Genet. 244:548-556(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambull R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansgore W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Macho R., Mueller M.,
RA Kreis M., Jelsky M., Puigdomenech P., Watson M., Schmidthein I.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Hancock I.,
RA Vos P., Hohenseil J., Zimmermann W., Wedter H., Radley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeker M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor F.,
RA Weitzenecker T., Bothe G., Hamspecker U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirksse W.,
RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doquett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
RA Pettitt A., Rajandream M.A., Lync M., Rones V., Rechmann S.,
RA Rorkova D., Bloeker H., Scharfe M., Grimm M., Loehert T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Danner D., Herzl A.,
RA Neumann S., Arquier A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaudi G., Mueundlein A., Feiber R.,

```

```

RA Schnabl S., Hillier R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon P., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
RA Frishman D., Haase D., Lemcke K., Hewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson K.K., de la Bastide M., Habermann K.,
RA Parnell I., Dedhia N., Gnoj I., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Sreco T., Kemp K.,
RA Kramer J., Fulton L., Hardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shether M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Granat S., Shohdy N., Hasegawa A., Hancock A., Lodhi M., Johnson A.,
RA Chen F., Marra M., Martienssen R., McCombie W.R.
RA *Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana.*
RN [4]
RP SEQUENCE OF 1-82 FROM N.A.
RC STRAIN=cv. Columbia;
RA Berthomieu P., Guerrier D., Giraudat J.
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED
CC AND ABNORMAL PROTEINS.
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN YOUNG STEMS, OLD
CC LEAVES. LOWEST LEVELS IN FLORAL BUDS, ANthers AND YOUNG LEAVES.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z14990; CAA78714.1;
DR EMBL: X72626; CAA51201.1;
DR EMBL: L00639; AAA32894.1;
DR EMBL: AL035524; CAB36765.1;
DR EMBL: AL161572; CAB79598.1;
DR EMBL: Z18473; CAA79198.1;
DR PIR: S32674; S32674.
DR HSP: P15731; IQCC.
DR InterPro: IPR000608; UBO_conjugat.
DR Pfam: PF00179; UQ_conj; 1.
DR ProDom: PD000461; UBO_conjugat; 1.
DR SMART: SM00212; UBCc; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubl conjugation pathway; Ligase; Multigene family.
FT BINDING 85 85 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 148 AA; 16553 MW; 17F247FC1801B8BD CRC64;
Query Match 16.78; Score 252; DB 1; Length 148;
Best Local Similarity 35.58; Prod No. 1.7e-13;
Matches 50; Conservative 26; Mismatches 51; Gaps 2;
QY 9 SOKALMLEKLSQEPVGVGRITLVDESIDYNNVEVAIFGLPNTLYEGYFKAHKFPIDY 68
DB 2 ASKRILKELKQKQDPPTSCSAGPVAE-DMFHQWATFMGPTDSDPYAGGVFLVTHRPDY 60

```

QY 69 PYSPTPEFLTKMHPNTEYMGDVCTSLHPVVDQPSGELPSERWNPNTQNVRTLLSVI 128
 DB 61 PKPKVFAERKTVHPNINSNGICLDL-----KEQSPALITVSKVLISCSL 107
 QY 129 SLGNENPITSPANDASVMER 149
 DB 108 SLGIDPNPDPLVPEIAHMYK 128

RESULT 34
 UHSA_HUMAN STANDARD; PRT: 147 AA.
 AC P51668;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 41, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Ubiquitin-conjugating enzyme E2-17 kDa 1 (EC 6.3.2.19) (Ubiquitin-
 protein ligase) (Ubiquitin carrier protein) (E2(17)KB 1).
 GN UBE2N1 OR UBCH5A OR UBCH5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94377440; PubMed=8090726;
 RA Scheffner M., Huibregtse J.M., Howley P.M.:
 RT "Identification of a human ubiquitin-conjugating enzyme that mediates
 the E6-AP-dependent ubiquitination of p53.";
 RJ Proc. Natl. Acad. Sci. U.S.A. 91:8797-8801(1994).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED AND
 CC ABNORMAL PROTEINS. FUNCTIONS IN THE E6/E6-AP-INDUCED
 CC UBIQUITINATION OF P53.
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X78140; CAN55019.1;
 CC HSSP: P15731; IYCC.
 CC Genew: HGNC:12474; UBE2D.
 CC MIM: 602961;
 CC InterPro: IPR000608; UBQ_conjugat.
 CC Pfam: PF00179; UBQ_conj. 1.
 CC ProDom: PD000461; UBQ_conjugat; 1.
 CC SMART: SM00212; UBQC; 1.
 CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 CC PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 CC KEGG: ubi conjugation pathway; ligase; Multigene family.
 CC FT BINDING 85 85 UBIQUITIN (BY SIMILARITY).
 CC SEQUENCE 147 AA: 16602 MW: 2596ED0179EE119D CRC64;

Query Match 16.6%; Score 251.5; DB 1; Length 147;
 Best Local Similarity 34.7%; Pred. No. 1.9e-13;
 Matches 52; Conservative 24; Mismatches 53; Indels 21; Gaps 3;
 QY 11 KALMIELKSLQEPVEGRITLVDSLDYNNVEAIFGLPNTLYEGGYFAHKEIPIDPY 70
 DB 4 KRIQKELSDLRDPPAKSAGPVG-DLFHQATIMGPPDSAYGGVFLTVHPTDYPF 62
 QY 71 SIPTPEFLTKMHPNTEYMGDVCTSLHPVVDQPSGELPSERWNPNTQNVRTLLSVI 130

DB 63 KPPIAFTTKIHPNINSNGICLDLIR-----SQSPALITVSKVLISCSL 109
 QY 131 LNEPNTSPANDASVMERKWRKSGDKKE 160
 DB 110 LCDPNPDDPLVPIAQIYK-----SDKE 132
 RESULT 35
 UBC8_ARATH STANDARD; PRT: 148 AA.
 AC P35131; Q42308; Q43276;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Ubiquitin-conjugating enzyme E2-17 kDa 8 (EC 6.3.2.19) (Ubiquitin-
 protein ligase 8) (Ubiquitin carrier protein 8) (UBCAT4A).
 GN UBC8 OR UBC4A OR AT5G41700 OK MWK23.24.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Columbia; TISSUE=leaf;
 RX MEDLINE=94035158; PubMed=8220461;
 RA Girod P.-A., Carpenter I.B., van Nocker S., Sullivan M.L.,
 RA Vierstra R.D.:
 RT "Homologs of the essential ubiquitin conjugating enzymes UBC1, 4, and
 RT 5 in yeast are encoded by a multigene family in Arabidopsis
 RT thaliana.";
 RJ Plant J. 3:545-552(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Columbia;
 RX MEDLINE=97471969; PubMed=9330910;
 RA Sato S., Kotani H., Nakamura Y., Kaneko I., Asamizu E., Fukami M.,
 RA Miyajima N., Tabata S.:
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
 RT features of the 1.6 Mb regions covered by twenty physically assigned
 RT pl clones.";
 RJ DNA Res. 4:215-230(1997).
 RN [3]
 RP SEQUENCE OF 1-74 FROM N.A.
 RX STRAIN=cv. Columbia;
 RA Bardet C., Dabos P., Tremoulaye D., Lescure B.:
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 36-148 FROM N.A.
 RX STRAIN=cv. Columbia; TISSUE=Green siliques;
 RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delseny M.:
 RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED
 CC AND ABNORMAL PROTEINS.
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN YOUNG STEMS, OLD
 CC LEAVES. LOWEST LEVELS IN FLORAL HDS, ANTHEMS AND YOUNG LEAVES.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

```

DR EMBL: Z14589; CAA78713.1;
DR EMBL: AR055233; BAH11476.1;
DR EMBL: Z37225; CAA85527.1;
DR EMBL: Z17692; CAA79036.1;
DR HSSP: Z15731; IQCC.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; UQ_conjugat; 1.
DR SMART: SM00122; UBCC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubi conjugation pathway; Ligase; Multigene family.
FT BINDING 85 85 UBIQUITIN (BY SIMILARITY).
FT CONFLICT 72 72 K -> M (IN REF. 3).
FT CONFLICT 80 80 S -> D (IN REF. 4).
FT CONFLICT 98 98 T -> P (IN REF. 4).
SQ SEQUENCE 148 AA; 16537 MW; 27CAAEABBBE74972 CRC64;

Query Match 16.6%; Score 251; DB 1; Length 148;
Best Local Similarity 35.5%; Pred. No. 2.1e-13;
Matches 50; Conservative 25; Mismatches 52; Indels 14; Gaps 2;

QY 9 SOKALMELKSLQEPVEGFRITLVDESDLYNNEVAIFGLPNTLYEGGYFKAHIKFPIDY 68
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 ASKRLKELKDLQDPPTSCSAGPVAE-DMEHQOATIMGPAESPYSGGVFLVTHFFPPDY 60
QY 69 PYSPTFERFLTKMHNPNIYENDVCISILHPPVDPPQSGELPSEKWNPTQNVRIILSVI 128
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 PFKPKVAFKVKVPHNINSNGSICLDIL-----KEOWSPALTISKVILSIC 107

QY 129 SLINPNTFSPANVDASVMFR 149
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
108 SLIDPNPDPLVPEIAHYK 128

RESULT 36
ID UBCB_XENLA STANDARD; PRT; 148 AA.
AC P35133;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 10 (EC 6.3.2.19) (Ubiquitin-
DE protein ligase 10) (Ubiquitin carrier protein 10).
GN UBC10 OR A15653300 OR K19EL10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Leaf;
RX MEDLINE=94035158; PubMed=8220461;
RA Girard P.-A., Carpenter I.B., van Nocker S., Sullivan M.L.,
RA Viorstra R.D.;
RT 5 in yeast are encoded by a multigene family in Arabidopsis
RT thaliana.*;
R3 Plant J. 3:545-552(1993).
RN 12;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=98403884; PubMed=9734815;
RA Kolani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT *Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones.*;
RL DNA Res. 5:203-216(1998);
RN 13;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

```

```

RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT Niren Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).*;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED
CC AND ABNORMAL PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC EMBL: Z14991; CAA78715.1;
DR EMBL: I00640; AAA32895.1;
DR EMBL: AB013388; BAB09792.1;
DR EMBL: AF326872; AAC41454.1;
DR PIR: S32672; S32672.
DR HSSP: P15731; IQCC.
DR InterPro: IPR000608; UQ_conjugat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; UQ_conjugat; 1.
DR SMART: SM00122; UBCC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubi conjugation pathway; Ligase; Multigene family.
FT BINDING 85 85 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 148 AA; 16537 MW; 2R83EDCIAD2AE657 CRC64;

Query Match 16.5%; Score 250; DB 1; Length 148;
Best Local Similarity 35.5%; Pred. No. 2.5e-13;
Matches 50; Conservative 25; Mismatches 52; Indels 14; Gaps 2;

QY 9 SOKALMELKSLQEPVEGFRITLVDESDLYNNEVAIFGLPNTLYEGGYFKAHIKFPIDY 68
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 ASKRLKELKDLQDPPTSCSAGPVAE-DMEHQOATIMGPSPSYAGGVFLVTHFFPPDY 60
QY 69 PYSPTFERFLTKMHNPNIYENDVCISILHPPVDPPQSGELPSEKWNPTQNVRIILSVI 128
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 PFKPKVAFKVKVPHNINSNGSICLDIL-----KEOWSPALTISKVILSIC 107

QY 129 SLINPNTFSPANVDASVMFR 149
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
108 SLIDPNPDPLVPEIAHYK 128

RESULT 37
ID UBCB_XENLA STANDARD; PRT; 179 AA.
AC P56616;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme X (EC 6.3.2.19) (Ubiquitin-protein
DE ligase) (Ubiquitin carrier protein).
GN UBC-X.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Fungiostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8155;
RN 1;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 63-80 AND 99-108.
RX MEDLINE=96298869; PubMed=8723350;

```


[illegible]

GenCore vers: on 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:33:10 : Search time 26.1765 Seconds
(without alignments)
1272.618 Million cell updates/sec

Title: us-09-930-026-l
Perfect score: 250
Sequence: 1 MALLAISLPETWIKTFEDK.....KGEASQSGSEGAQGLAPS 250

Scoring table: GSI06

Gapop: 60.0, Gapext: 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 507025

Minimum DB seq length: 0

Maximum DB seq length: 70

Post-processing: Listing first 135 summaries

Database: A_geneseq_101002.*

```

1: /SID52/gcdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	2.8	9	20	BCR-Ab1 epitope (a
2	7	2.8	9	20	AAY3352
3	7	2.8	9	20	AAY40174
4	7	2.8	9	20	AAY26692
5	7	2.8	16	23	BCR-Ab1-derived li
6	7	2.8	17	23	ABL-i peptide. Un
7	7	2.8	30	9	ABL-i peptide. Un
8	7	2.8	30	11	peptide encoded by
9	7	2.8	30	11	Signal peptide der
10	7	2.8	50	23	Human ORFX protein
11	7	2.8	50	23	Human ORFX protein
12	7	2.8	50	23	Human ORFX protein
13	7	2.8	50	23	Human ORFX protein
14	7	2.8	50	23	Human ORFX protein
15	7	2.8	50	23	Human ORFX protein
16	7	2.8	50	23	Human ORFX protein
17	7	2.8	50	23	Human ORFX protein
18	7	2.8	50	23	Human ORFX protein
19	7	2.8	50	23	Human ORFX protein
20	7	2.8	50	23	Human ORFX protein
21	7	2.8	50	23	Human ORFX protein
22	7	2.8	50	23	Human ORFX protein
23	7	2.8	50	23	Human ORFX protein
24	7	2.8	50	23	Human ORFX protein
25	7	2.8	50	23	Human ORFX protein
26	7	2.8	50	23	Human ORFX protein
27	7	2.8	50	23	Human ORFX protein
28	7	2.8	50	23	Human ORFX protein
29	7	2.8	50	23	Human ORFX protein
30	7	2.8	50	23	Human ORFX protein
31	7	2.8	50	23	Human ORFX protein
32	7	2.8	50	23	Human ORFX protein
33	7	2.8	50	23	Human ORFX protein
34	7	2.8	50	23	Human ORFX protein
35	7	2.8	50	23	Human ORFX protein
36	7	2.8	50	23	Human ORFX protein
37	7	2.8	50	23	Human ORFX protein
38	7	2.8	50	23	Human ORFX protein
39	7	2.8	50	23	Human ORFX protein
40	7	2.8	50	23	Human ORFX protein
41	7	2.8	50	23	Human ORFX protein
42	7	2.8	50	23	Human ORFX protein
43	7	2.8	50	23	Human ORFX protein
44	7	2.8	50	23	Human ORFX protein
45	7	2.8	50	23	Human ORFX protein
46	7	2.8	50	23	Human ORFX protein
47	7	2.8	50	23	Human ORFX protein
48	7	2.8	50	23	Human ORFX protein
49	7	2.8	50	23	Human ORFX protein
50	7	2.8	50	23	Human ORFX protein
51	7	2.8	50	23	Human ORFX protein
52	7	2.8	50	23	Human ORFX protein
53	7	2.8	50	23	Human ORFX protein
54	7	2.8	50	23	Human ORFX protein
55	7	2.8	50	23	Human ORFX protein
56	7	2.8	50	23	Human ORFX protein
57	7	2.8	50	23	Human ORFX protein
58	7	2.8	50	23	Human ORFX protein
59	7	2.8	50	23	Human ORFX protein
60	7	2.8	50	23	Human ORFX protein
61	7	2.8	50	23	Human ORFX protein
62	7	2.8	50	23	Human ORFX protein
63	7	2.8	50	23	Human ORFX protein
64	7	2.8	50	23	Human ORFX protein
65	7	2.8	50	23	Human ORFX protein
66	7	2.8	50	23	Human ORFX protein
67	7	2.8	50	23	Human ORFX protein
68	7	2.8	50	23	Human ORFX protein
69	7	2.8	50	23	Human ORFX protein
70	7	2.8	50	23	Human ORFX protein
71	7	2.8	50	23	Human ORFX protein
72	7	2.8	50	23	Human ORFX protein
73	7	2.8	50	23	Human ORFX protein
74	7	2.8	50	23	Human ORFX protein
75	7	2.8	50	23	Human ORFX protein
76	7	2.8	50	23	Human ORFX protein
77	7	2.8	50	23	Human ORFX protein
78	7	2.8	50	23	Human ORFX protein
79	7	2.8	50	23	Human ORFX protein
80	7	2.8	50	23	Human ORFX protein
81	7	2.8	50	23	Human ORFX protein
82	7	2.8	50	23	Human ORFX protein
83	7	2.8	50	23	Human ORFX protein
84	7	2.8	50	23	Human ORFX protein
85	7	2.8	50	23	Human ORFX protein
86	7	2.8	50	23	Human ORFX protein
87	7	2.8	50	23	Human ORFX protein
88	7	2.8	50	23	Human ORFX protein
89	7	2.8	50	23	Human ORFX protein
90	7	2.8	50	23	Human ORFX protein
91	7	2.8	50	23	Human ORFX protein
92	7	2.8	50	23	Human ORFX protein
93	7	2.8	50	23	Human ORFX protein
94	7	2.8	50	23	Human ORFX protein
95	7	2.8	50	23	Human ORFX protein
96	7	2.8	50	23	Human ORFX protein
97	7	2.8	50	23	Human ORFX protein
98	7	2.8	50	23	Human ORFX protein
99	7	2.8	50	23	Human ORFX protein
100	7	2.8	50	23	Human ORFX protein

Core peptide #40 f
Peptide derived fr
Core peptide #7 fo
Synthetic CD4 anal
Peptide derived fr
Peptide derived fr
SH3-binding prolin
SH3-binding prolin
Human peptide #207
CDR3-like constrai
TEDPK-8 derivative
Placental lactogen
Membrane occupatio
1 domain peptide p
Mammalian CD4 T1-1
IgE Fcεpsilon RI b
IgE Fcεpsilon RI b
IgE Fcεpsilon RI b
granulocyte-colony
HIV gp120 protein
HIV gp120 protein
Sulfolobus solfata
Eubacterial MatS1
Amino acid sequenc
Peptide #4993 enco
Novel human diagno
Human brain expres
Human peptide enco
Streptococcus poly
Human protein SEQ
Mouse type II inte
Human secreted pro
Drosophila me'lanog
Human secreted pro
Human colon cancer
Human polypeptide
Propionibacterium
Human gαuc 34-enco
Human αlbumin fusi
Novel human diagno
Human nervous syst
Novel human diagno
Human nervous syst
Human cadherin-lik
Human immune/haema
Human endoprotease
Propionibacterium
Staphylococcus aur
Novel human diagno
Peptide #8403 enco
Human brain expres
Human bone marrow
Peptide #8704 enco
Human peptide enco
Human ORF1543 prot
Human immune/haema
Peptide #11877 enc
Protein #5227 enco
Human brain expres
Human bone marrow
Human immune/haema
Peptide #8427 enco
Peptide #12377 enc
Human peptide enco
Lactococcus lactis
Human testicular a
Propionibacterium
Human nervous syst
Human reproductive
Human ORF2239 prot
Peptide #2151 enco

84 6 2.4 65 22 AAM55470 Human brain expro
85 6 2.4 65 22 AAM67856 Human bone marrow
86 6 2.4 65 22 AAM15677 Peptide #2111 enco
87 6 2.4 65 22 AAM28180 Peptide #2217 enco
88 6 2.4 65 22 AAG37402 Human peptide enco
89 6 2.4 66 22 AAU49861 Propionibacterium
90 6 2.4 66 22 AAC02956 Human polypeptide
91 6 2.4 67 23 ABP32182 Human ORF155 prot
92 6 2.4 69 22 AAJ86564 Novel human connec
93 6 2.4 69 22 AAU42480 Propionibacterium
94 6 2.4 69 22 AAC04738 Human polypeptide
95 6 2.4 70 22 ABF19607 Novel human diagno
96 6 2.4 70 23 ABP40129 Staphylococcus epi
97 5 2.0 5 18 AAW10247 Peptide used to pr
98 5 2.0 5 18 AAW11686 Core peptide #36 f
99 5 2.0 5 20 AAW87680 Peptide derived fr
100 5 2.0 5 20 AAW84089 Human TSG-6 bindin
101 5 2.0 5 22 AAF02363 TSG-6 related pept
102 5 2.0 6 18 AAW11687 Core peptide #37 f
103 5 2.0 6 20 AAW87681 Peptide derived fr
104 5 2.0 6 23 AAU80049 Human RalBpi relat
105 5 2.0 7 34 AAR42272 TGF-beta1 inhibito
106 5 2.0 7 18 AAW11688 Core peptide #38 f
107 5 2.0 7 20 AAY17026 Heat shock protein
108 5 2.0 7 20 AAW87682 Peptide derived fr
109 5 2.0 7 21 AAB10669 L. mesenteroides a
110 5 2.0 7 22 AAU72168 Melanoma antigen,
111 5 2.0 7 23 ABP48416 Zinc finger protei
112 5 2.0 7 23 ABP48417 Zinc finger protei
113 5 2.0 7 23 ABP48493 Zinc finger protei
114 5 2.0 8 14 AAR36094 Javelin peptide #1
115 5 2.0 8 14 AAR36095 Hepatitis C virus
116 5 2.0 8 14 AAR36096 Hepatitis C virus
117 5 2.0 8 15 AAR72181 Hepatitis C virus
118 5 2.0 8 15 AAR72182 Hepatitis C virus
119 5 2.0 9 15 AAR72183 Hepatitis C virus
120 5 2.0 9 15 AAR72184 Hepatitis C virus
121 5 2.0 8 16 AAR83814 SH3-binding prolin
122 5 2.0 8 16 AAR56492 Zea mays soluble s
123 5 2.0 8 20 AAY16855 Heat shock protein
124 5 2.0 8 21 AAH05932 Test antigen #8 to
125 5 2.0 8 22 AAU69074 Bacterial conserve
126 5 2.0 8 23 AHC67334 Human ADP1 tryptic
127 5 2.0 9 15 AAR47632 Hepatitis C virus
128 5 2.0 9 16 AAR83815 SH3-binding prolin
129 5 2.0 9 16 AAR83817 SH3-binding prolin
130 5 2.0 9 16 AAR83818 SH3-binding prolin
131 5 2.0 9 16 AAR83819 SH3-binding prolin
132 5 2.0 9 16 AAR83820 SH3-binding prolin
133 5 2.0 9 16 AAR83821 SH3-binding prolin
134 5 2.0 9 17 AAW49322 Human leucocyte an

ALIGNMENTS

RESULT 1

AAV53352
ID AAY53352 standard; Protein; 9 AA.

XX
AC AAY53352;

DT 18-JAN-2000 (first entry)

DE Her-Abl epitope (aa 1499-1509) binds H1A-A3/A11.

XX
KW lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;
KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;
KW human immunodeficiency virus; hepatitis B virus; papilloma virus;
XX
XX melanoma; malaria; parasite.
OS Synthetic.

OS Homo sapiens.
XX
PN FR2776926-A1.
XX
PD 08-OCT-1999.
XX
PF 07-APR-1998; 98FR-0004323.
XX
PR 07-APR-1998; 98FR-0004323.
XX (INRM) INERM INST NAT SANTE & RECH MEDICALE.
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (INSP) INST PASTEUR LILLE.
XX
PI Le Gal, FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;
XX
DR WPI: 1999-583113/50.
XX
PT New lipopeptide containing lipid regions and two epitopes, all
PT separated by peptide spacers that impart hydrophilicity, useful in
PT vaccines
XX
PS Disclosure; Page 17: 35pp; French.
XX
CC The invention relates to the generation of a lipopeptide comprising at
CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)
CC epitope and at least one lipid residue with (i) the epitopes and lipid
CC portion and (ii) the epitopes, being separated independently by peptide
CC spacers. These spacers comprise sequences of amino acids which carry an
CC overall electrical charge in neutral media to ensure that the
CC lipopeptide is hydrophilic. The peptides AAY53301-Y53549 represents
CC examples of peptide epitopes used to generate the lipopeptides. These are
CC used in therapeutic or prophylactic compositions and vaccines to induce
CC specific immune responses against human immunodeficiency, hepatitis B or
CC papilloma viruses; p53 of melanoma or the malaria parasite.
XX
SQ Sequence 9 AA;
Query Match 2.8%; Score 7; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 23 LFSALIK 29
DB 2 LFSALIK 8
RFSULT 2
AAV40174
ID AAY40174 standard; Reptide; 9 AA.
XX
AC AAY40174;
XX
DT 19-NOV-1999 (first entry)
XX
DE Amino acid sequence of a chronic myeloid leukaemia epitope.
XX
KW Cytotoxic T cell; T lymphocyte; CD8+ epitope; T helper cell;
KW CD4+ epitope; B epitope; lipopeptide; interferon gamma; adjuvant;
KW vaccine; tumor; infection; immune response; cytokine profile;
KW acquired immune deficiency syndrome; papilloma; cancer; hepatitis;
KW autoimmune disease.
OS Homo sapiens.
XX
PN FR2774687-A1.
XX
PD 13-AUG-1999.
XX
PF 06-FEB-1998; 98FR-0001439.
XX
PR 06-FEB-1998; 98FR-0001439.
XX

PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR ILLUE.
 XX Thiam K, Guillet JG, Ver Waerde C, Auriault C, Gras MH, Loing E;
 XX WPI; 1999-5:0734/43.
 XX New lipopeptide confiting C-terminal interferon-gamma fragment with
 PT attached lipophilic groups, used as interferon mimic, e.g. for treating
 PT cancer or virus infection
 XX Disclosure: Page 33; 53pp; French.
 XX AAY0123-V40379 represent epitopes that are able to activate cytotoxic
 CC T lymphocytes (CD8+ epitopes). T helper cells (CD4+ epitopes), or
 CC B epitopes recognized by corresponding antibodies. The epitopes may be
 CC used in the composition of the invention. The specification describes a
 CC lipopeptide that has a peptide part derived from mammalian interferon
 CC gamma (IFN γ) and one or more lipophilic parts comprising a linear or
 CC branched, (un)saturated 4-20C hydrocarbyl chain or a steroid. The
 CC lipopeptide mimics the activity of IFN γ . Compositions comprising the
 CC lipopeptide are used to treat or prevent any condition that responds
 CC to IFN γ , and as adjuvant for vaccines (particularly those directed
 CC against tumors, viral or parasitic infections), to stimulate or
 CC (re)orient the immune response between types 1 and 2 cytokine profiles.
 CC Particular applications are treatment of infections (particularly
 CC viral, e.g. acquired immune deficiency syndrome, papilloma (cancer) and
 CC hepatitis, but also bacterial, fungal, parasitic or helminth); cancers
 CC (particularly of kidney, cutaneous T cells or ovary, chronic
 CC myelogenous leukemia or mesothelioma), allergy; and autoimmune
 CC diseases.
 XX Sequence 9 AA:
 SQ Query Match 2.8%; Score 7; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LFSALIK 29
 DB 2 LFSALIK 8

RESULT 3
 AAY26690
 ID MAY26690 standard; peptide; 9 AA.
 XX AC AAY26690;
 XX 14 SEP-1999 (first entry)
 XX BCR-ABL-derived lipopeptide epitope (aa1499-1509) for mixed micelles.
 XX Micelle: microaggregate; induction: immune response; lipopeptide; CTL;
 KW cytotoxic T-lymphocyte; epitope; lipid; helper T-lymphocyte; HPL; HBV;
 KW tetanus; toxin; vaccine; HIV; Hepatitis B virus; papilloma virus; p53;
 KW melanoma; Plasmodium falciparum; malaria.
 XX Synthetic.
 OS Avicel murine leukemia virus.
 XX PR2771640-AL.
 XX 04 JUN-1999.
 XX 03-DEC-1997; 97FR-C015246.
 XX 03-DEC-1997; 97FR-C015246.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR ILLUE.
 XX

PI Bossus M, Bourgault VI, Gras-Masse H, Guillet JG, Lippens G;
 PI Tartar A, Wieruszski JM;
 XX WPI; 1999-349509/30.
 XX Immunogenic lipopeptide micelles - comprising lipopeptides
 PT containing cytotoxic and helper T-lymphocyte epitopes
 PT Disclosure: Page 30; 60pp; French.
 XX The invention relates to the generation of mixed micelles or
 CC microaggregates for inducing an immune response comprise: (a) a first
 CC lipopeptide comprising at least one CTL (cytotoxic T-lymphocyte) epitope
 CC and at least one lipid unit; and (b) a second lipopeptide comprising at
 CC least one HPL (helper T-lymphocyte) epitope and at least one lipid unit
 CC different from that of the first lipopeptide. This peptide represents
 CC an example of a lipopeptide epitope used in the invention and is derived
 CC from the BCR protein of the Abelson murine leukemia virus. The
 CC immunogenic lipopeptide micelles are used in vaccines, especially against
 CC HIV, hepatitis B virus (HBV), papilloma viruses, p53, melanoma or
 CC Plasmodium falciparum malaria.
 XX Sequence 9 AA:
 SQ Query Match 2.8%; Score 7; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LFSALIK 29
 DB 2 LFSALIK 8

RESULT 4
 AAU78972
 ID AAU78972 standard; Peptide; 16 AA.
 XX AC AAU78972;
 XX 18-JUN-2002 (first entry)
 XX ABL-1 peptide.
 XX Nuclear localisation signal; NLS; protein delivery;
 KW fusion protein; membrane penetrating peptide; ABL-1;
 KW protein transduction domain.
 XX Unidentified.
 OS WO200218572-A2.
 XX 07-MAR-2002.
 XX 23-AUG-2001; 2001WO-US26421.
 XX 25-AUG-2000; 2000US-227647P.
 XX 07-FEB-2001; 2001GB-0003110.
 XX (AVEI) AVENTIS PHARM INC.
 XX Guo Y, Morse CC, Yao Z, Keesler GA;
 XX WPI; 2002-304256/34.
 XX New fusion proteins comprising membrane penetrating peptides, useful as
 PT in vivo, ex vivo or in vitro intracellular carriers or delivery devices
 PT for a compound of interest (e.g. peptide, protein, chemical entity,
 PT nucleic acid)
 XX Example 5; Fig 5; 45pp; English.
 XX This invention relates to a novel fusion protein, which comprises a
 CC membrane penetrating peptide attached to a compound of interest.

CC The membrane penetrating peptide of the fusion protein is derived from a
 CC nuclear localisation signal and may be the nuclear localisation signal
 CC from human period protein hPER1. The fusion protein is useful for
 CC delivery of a compound of interest into a cell. The fusion protein is
 CC useful as in vivo, ex vivo or in vitro intracellular delivery devices
 CC for a compound of interest (e.g. peptide, protein, chemical entity,
 CC nucleic acid). In particular, the polypeptides are useful as protein
 CC carriers for delivery of compounds to cells. The present sequence
 CC represents the ABL-1 peptide sequence which was identified by searching
 CC the NCBI protein database with a degenerative algorithm of the
 CC invention. This sequence was then shown to translocate to the nucleus in
 CC the examples of the invention.

XX
 SQ Sequence 16 AA;

Query Match 2.8%; Score 7; DB 23; Length 16;
 Best Local Similarity 100.0%; Prod. No. 15;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 LFSALIK 29

 DB 5 LFSALIK 11

RESULT 5
 AAL78940
 ID AAL78940 standard; Peptide: 17 AA.

XX AC AAL78940;

XX DT 18-JUN-2002 (first entry)

DE ABL-1 peptide.

XX KW Nuclear localisation signal; NLS; protein; delivery; ABL-1;
 KW fusion protein; membrane penetrating peptide.

XX Unidentified.

XX PN WO200218572-A2.

XX PD 07-MAR-2002.

XX PF 23-AUG-2001; 2001WG-US26421.

XX PR 25-AUG-2000; 2000US-227647P.

XX PR 07-FEB-2001; 2001GB-0003110.

XX PA (AVPT) AVENTIS PHARM INC.

XX PI Guo Y, Morse CC, Yao Z, Keesler GA;

XX LR WPI: 2002-304256/34.

XX PI New fusion proteins comprising membrane penetrating peptides, useful as
 PI in vivo, ex vivo or in vitro intracellular carriers or delivery devices
 PI for a compound of interest (e.g. peptide, protein, chemical entity,
 PT nucleic acid).

XX Example 2; Page 28; 45pp; English.

XX This invention relates to a novel fusion protein, which comprises a
 CC membrane penetrating peptide attached to a compound of interest.
 CC The membrane penetrating peptide of the fusion protein is derived from a
 CC nuclear localisation signal and may be the nuclear localisation signal
 CC from human period protein hPER1. The fusion protein is useful for
 CC delivery of a compound of interest into a cell. The fusion protein is
 CC useful as in vivo, ex vivo or in vitro intracellular delivery devices
 CC for a compound of interest (e.g. peptide, protein, chemical entity,
 CC nucleic acid). In particular, the polypeptides are useful as protein
 CC carriers for delivery of compounds to cells. The present sequence
 CC represents the human ABL-1 peptide used in an assay to analyse the
 CC ability of different peptides to penetrate cellular membranes in the

CC examples of the invention.

XX
 SQ Sequence 17 AA;

Query Match 2.8%; Score 7; DB 23; Length 17;
 Best Local Similarity 100.0%; Prod. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 LFSALIK 29

.

DB 6 LFSALIK 12

RESULT 6

AAP80653

ID AAP80653 standard; protein; 30 AA.

XX AC AAP80653;

XX DT 08-OCT-1990 (first entry)

DE Peptide encoded by promoter and signal peptide sequence.

XX KW Tumour necrosis factor; anticancer drug; promoter; signal peptide.

XX PN JP63258583-A.

XX PD 26-DEC-1988.

XX PF 23-DEC-1986; 86JP-0307215.

XX PR 23-DEC-1986; 86JP-0307215.

XX PR 17-JUN-1987; 87JP-0152353.

XX PA (RIKA) RIKAGAKU KENKYUSHO.

XX DR WPI: 1988-348693/49.

XX DR N-PSDB: AAN81304.

XX PT New plasmid, for anti-cancer drug prodn.
 PT contg. DNA fragments with regions encoding antitumour
 PT polypeptide and accelerating extra-cellular secretion to host cells.

XX PS Disclosure; gc 506; 30pp; Japanese.

XX CC The peptide is expressed from the promoter and signal peptide sequence
 CC for prodn. of TNF. These are inserted into a plasmid with the KIL gene
 CC from plasmid pMB9, which accelerates the extracellular secretion of TNF,
 CC and the ex-promoter.

XX SQ Sequence 30 AA;

Query Match 2.8%; Score 7; DB 9; Length 30;
 Best Local Similarity 100.0%; Prod. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 CVSLIGT 85

.

DB 12 CVSLIGT 18

RESULT 7

AAR05809

ID AAR05809 standard; protein; 30 AA.

XX AC AAR05809;

XX DT 14-NOV-1990 (first entry)

DE Signal peptide derived from alkophilic Bacillus strain 170.

XX KW Antitumour; cancer; pBXTNF9; E.coli; ds.

XX

QS Synthetic.
 PN JF02145196-A.
 XX
 PD 04-JUN-1990.
 XX
 PP 28-NOV-1988; 88JP-0298211.
 XX
 PR 28-NOV-1988; 88JP-0298211.
 XX
 PA (TELJ) TELJIN KK.
 XX
 DR WPI: 1990-214248/28.
 DR N-PSDB: AAQ05181.
 XX
 PT New plasmid -
 contains DNA domains encoding polypeptide with antitumour activity.
 XX
 PS Disclosure; : P; Japanese.
 XX
 CC Plasmid pBX:INF may be expressed by transformed E.coli HB101 cells
 CC coupled with promoter from the first domain of chromosomal DNA of
 CC alkalophilic, Bacillus No.170.
 XX
 SQ Sequence 30 AA:
 Query Match 2.8%; Score 7; DB 11; Length 30;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 79 CVSLLET 85
 DE 12 CVSLLET 18
 |||||
 RESULT 8
 ABP01597
 ID ABP01597 standard; Protein: 50 AA.
 XX
 AC ABP01597;
 XX
 DT 25-JUN-2002 (first entry)
 XX
 DE Human ORFX protein sequence SEQ ID NO:3176.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 OS Homo sapiens.
 XX
 PN WG200192523-A2.
 XX
 PD 05-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US10836.
 XX
 PR 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach ME;
 XX
 DR WPI: 2002-106308/14.
 DR N-PSDB: ABN17349.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders
 XX
 PS Disclosure; SEQ ID 3176; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification)). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 50 AA:
 Query Match 2.8%; Score 7; DB 23; Length 50;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 79 CVSLLET 85
 DE 17 CVSLLET 23
 |||||
 RESULT 9
 AAW11689
 ID AAW11689 standard; peptide; 6 AA.
 XX
 AC AAW11689;
 XX
 DT 17-OCT-1997 (first entry)
 XX
 DE Core peptide #39 for synthetic CD4 analogue.
 XX
 KW CD4; therapy; macrocyclic peptidomimetic; inhibitor; MHC class II;
 KW helper T-cell; co-receptor; antigen receptor; immune response; antigen;
 KW signal transduction pathway; human; autoimmune disease; graft rejection;
 KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus.
 XX
 OS Synthetic.
 XX
 PN WO9701350-A1.
 XX
 PD 16-JAN-1997.
 XX
 PF 28-JUN-1996; 96WO-US11176.
 XX
 PR 20-SEP-1995; 95US-0004034.
 PR 29-JUN-1995; 95US-0000710.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Huang Z, Korngold R;
 XX
 DR WPI: 1997-100011/09.

XX New peptide(s) based on CD4 sequences - which suppress CD4 T-cell
 PT mediated responses, partic. for treating auto-immune disease or
 PT preventing graft rejection.
 XX
 PS Claim 1: Page 42; 60pp; English.
 XX
 CC AAW11651-W11694 represent core peptides that can be used in the
 CC synthetic CD4 peptides of the invention (see AAW11642-W11650). These core
 CC peptides can also be used in macrocyclic peptidomimetics of the
 CC invention. In the peptidomimetics, the core peptide is linked to a
 CC 10-member, heterocyclic ring. The peptides and peptidomimetics mimic the
 CC surface of CD4 molecules and interfere with the interaction of CD4 and
 CC MHC class II, gene products. CD4 molecules are glycoproteins expressed on
 CC the surface of helper T-cells. CD4 molecules are co-receptors for the
 CC T-cell antigen receptor, and are implicated in the signal transduction
 CC pathway for activation of helper T-cells. In the signal transduction
 CC pathway, the presence of an antigen leads to the activation of an
 CC antigen-specific helper T-cell, through the action of the pathway.
 CC undesired immune responses in a human subject (such as responses against
 CC "self-antigens") cause autoimmune diseases in humans. The peptides and
 CC peptidomimetics can therefore be used for suppressing a human CD4 T-cell
 CC mediated response. They can be used to treat autoimmune diseases such as
 CC rheumatoid arthritis, multiple sclerosis and systemic lupus
 CC erythematosus. They can also be used to prevent graft rejection.
 XX
 SQ Sequence 6 AA:
 Query Match 2.48; Score 6; DR 18; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e-05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 216 LSDSGQ 221
 Db 1 LSDSGQ 6
 RESULT 1:
 AAW87683
 ID AAW87683 standard; peptide: 6 AA.
 XX
 AC AAW87683;
 XX
 DT 04-MAR-1999 (first entry)
 XX
 DE Peptide derived from D4 domain FG ridge of human CD4.
 XX
 DE Human CD4; cyclic peptide; suppression; allograft rejection;
 KW CD4-dependent T-cell mediated immune response; multiple sclerosis.
 KW
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX US5846933-A.
 XX
 PD 08-DIC-1996.
 XX
 XX 28-JUN-1996; 96US-0672610.
 XX
 XX 28-JUN-1996; 96US-0672610.
 XX
 XX (HUANG/) HUANG Z.
 PA (KORN/) KORNIGOLD R.
 XX
 XX Huang Z, Kornigold R;
 XX
 XX WPI: 1999-059113/05.
 XX
 XX Inhibition of T cell immune response in vivo - by administering
 PT cyclic CD4-derived peptide(s)
 XX
 XX Disclosure: Columns 9-10; 27pp; English.
 PS
 XX

CC The present peptide is derived from human CD4. It is the core peptide of
 CC a cyclic peptide that is used in the method of the invention, which is
 CC used for suppressing a human CD4-dependent T-cell-mediated immune
 CC response. The method comprises administering a peptide of formula:
 CC R-X-Z-A1-A2-A3-A4-A5-Z'-X'-R, or
 CC R-X-Z-(CORE PEPTIDE)-Z'-X'-R,
 CC where R=H or acetyl; R'=OH or NH2; A1=an L-amino acid selected
 CC from N, Q and D; A2=an L-amino acid selected from S, T, D and N;
 CC A3=an L-amino acid selected from N, D, E and Q; A4=an L-amino acid
 CC selected from Q, N, E and M; A5=an L-amino acid selected from I, L, V
 CC and A; X=an L-amino acid or a di- or tripeptide of D- or L-amino acids
 CC selected from Y, W and F, provided that no amino acid is selected more
 CC than once, or X is absent; X'=a D- or L-amino acid selected from Y,
 CC W, F, I and L or a dipeptide of D- or L-amino acids selected from I, and
 CC I, or X' is absent; Z and Z' are amino acids that are linked to each
 CC other so that the peptide is cyclic; and CORE PEPTIDE is a sequence
 CC selected from AAW87645-84. The method is used especially for treating
 CC allograft rejection or multiple sclerosis.
 XX
 SQ Sequence 6 AA:
 Query Match 2.48; Score 6; DR 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e-05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 216 LSDSGQ 221
 Db 1 LSDSGQ 6
 RESULT 1:
 AAW11690
 ID AAW11690 standard; peptide: 7 AA.
 XX
 AC AAW11690;
 XX
 DT 17-OCT-1997 (first entry)
 XX
 DE Core peptide #40 for synthetic CD4 analogue.
 XX
 KW CD4; therapy; macrocyclic peptidomimetic; inhibitor; MHC class II;
 KW helper T-cell; co-receptor; antigen receptor; immune response; antigen;
 KW signal transduction pathway; human; autoimmune disease; graft rejection;
 KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus.
 XX
 OS Synthetic.
 OS
 XX W09701350-A1.
 XX
 XX 16-JAN-1997.
 XX
 XX 28-JUN-1996; 96WO-US11176.
 XX
 XX 20-SEP-1995; 95US-0004034.
 XX
 XX 29-JUN-1995; 95US-0000710.
 XX
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 XX Huang Z, Kornigold R;
 XX
 XX WPI: 1997-100011/09.
 XX
 XX New peptide(s) based on CD4 sequences - which suppress CD4 T-cell
 PT mediated responses, partic. for treating auto-immune disease or
 PT preventing graft rejection.
 XX
 XX Claim 1: Page 43; 66pp; English.
 XX
 CC AAW11651-W11694 represent core peptides that can be used in the
 CC synthetic CD4 peptides of the invention (see AAW11642-W11650). These core
 CC peptides can also be used in macrocyclic peptidomimetics of the
 CC invention. In the peptidomimetics, the core peptide is linked to a
 CC 10-member, heterocyclic ring. The peptides and peptidomimetics mimic the
 CC surface of CD4 molecules and interfere with the interaction of CD4 and
 CC MHC class II, gene products. CD4 molecules are glycoproteins expressed on
 CC the surface of helper T-cells. CD4 molecules are co-receptors for the
 CC T-cell antigen receptor, and are implicated in the signal transduction
 CC pathway for activation of helper T-cells. In the signal transduction
 CC pathway, the presence of an antigen leads to the activation of an
 CC antigen-specific helper T-cell, through the action of the pathway.
 CC undesired immune responses in a human subject (such as responses against
 CC "self-antigens") cause autoimmune diseases in humans. The peptides and
 CC peptidomimetics can therefore be used for suppressing a human CD4 T-cell
 CC mediated response. They can be used to treat autoimmune diseases such as
 CC rheumatoid arthritis, multiple sclerosis and systemic lupus
 CC erythematosus. They can also be used to prevent graft rejection.
 XX
 SQ Sequence 6 AA:
 Query Match 2.48; Score 6; DR 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e-05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 216 LSDSGQ 221
 Db 1 LSDSGQ 6
 RESULT 1:
 AAW11690
 ID AAW11690 standard; peptide: 7 AA.
 XX
 AC AAW11690;
 XX
 DT 17-OCT-1997 (first entry)
 XX
 DE Core peptide #40 for synthetic CD4 analogue.
 XX
 KW CD4; therapy; macrocyclic peptidomimetic; inhibitor; MHC class II;
 KW helper T-cell; co-receptor; antigen receptor; immune response; antigen;
 KW signal transduction pathway; human; autoimmune disease; graft rejection;
 KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus.
 XX
 OS Synthetic.
 OS
 XX W09701350-A1.
 XX
 XX 16-JAN-1997.
 XX
 XX 28-JUN-1996; 96WO-US11176.
 XX
 XX 20-SEP-1995; 95US-0004034.
 XX
 XX 29-JUN-1995; 95US-0000710.
 XX
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 XX Huang Z, Kornigold R;
 XX
 XX WPI: 1997-100011/09.
 XX
 XX New peptide(s) based on CD4 sequences - which suppress CD4 T-cell
 PT mediated responses, partic. for treating auto-immune disease or
 PT preventing graft rejection.
 XX
 XX Claim 1: Page 43; 66pp; English.
 XX
 CC AAW11651-W11694 represent core peptides that can be used in the
 CC synthetic CD4 peptides of the invention (see AAW11642-W11650). These core
 CC peptides can also be used in macrocyclic peptidomimetics of the
 CC invention. In the peptidomimetics, the core peptide is linked to a
 CC 10-member, heterocyclic ring. The peptides and peptidomimetics mimic the
 CC surface of CD4 molecules and interfere with the interaction of CD4 and
 CC MHC class II, gene products. CD4 molecules are glycoproteins expressed on
 CC the surface of helper T-cells. CD4 molecules are co-receptors for the
 CC T-cell antigen receptor, and are implicated in the signal transduction
 CC pathway for activation of helper T-cells. In the signal transduction
 CC pathway, the presence of an antigen leads to the activation of an
 CC antigen-specific helper T-cell, through the action of the pathway.
 CC undesired immune responses in a human subject (such as responses against
 CC "self-antigens") cause autoimmune diseases in humans. The peptides and
 CC peptidomimetics can therefore be used for suppressing a human CD4 T-cell
 CC mediated response. They can be used to treat autoimmune diseases such as
 CC rheumatoid arthritis, multiple sclerosis and systemic lupus
 CC erythematosus. They can also be used to prevent graft rejection.
 XX
 SQ Sequence 6 AA:
 Query Match 2.48; Score 6; DR 20; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e-05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 216 LSDSGQ 221
 Db 1 LSDSGQ 6
 RESULT 1:
 AAW11690
 ID AAW11690 standard; peptide: 7 AA.
 XX
 AC AAW11690;
 XX
 DT 17-OCT-1997 (first entry)
 XX
 DE Core peptide #40 for synthetic CD4 analogue.
 XX
 KW CD4; therapy; macrocyclic peptidomimetic; inhibitor; MHC class II;
 KW helper T-cell; co-receptor; antigen receptor; immune response; antigen;
 KW signal transduction pathway; human; autoimmune disease; graft rejection;
 KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus.
 XX
 OS Synthetic.
 OS
 XX W09701350-A1.
 XX
 XX 16-JAN-1997.
 XX
 XX 28-JUN-1996; 96WO-US11176.
 XX
 XX 20-SEP-1995; 95US-0004034.
 XX
 XX 29-JUN-1995; 95US-0000710.
 XX
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 XX Huang Z, Kornigold R;
 XX
 XX WPI: 1997-100011/09.
 XX
 XX New peptide(s) based on CD4 sequences - which suppress CD4 T-cell
 PT mediated responses, partic. for treating auto-immune disease or
 PT preventing graft rejection.
 XX
 XX Claim 1: Page 43; 66pp; English.
 XX
 CC AAW11651-W11694 represent core peptides that can be used in the
 CC synthetic CD4 peptides of the invention (see AAW11642-W11650). These core
 CC peptides can also be used in macrocyclic peptidomimetics of the
 CC invention. In the peptidomimetics, the core peptide is linked to a
 CC 10-member, heterocyclic ring. The peptides and peptidomimetics mimic the

CC surface of CD4 molecules and interfere with the interaction of CD4 and
 CC MHC class II, gene products. CD4 molecules are glycoproteins expressed on
 CC the surface of helper T-cells. CD4 molecules are co-receptors for the
 CC T-cell antigen receptor, and are implicated in the signal transduction
 CC pathway for activation of helper T-cells. In the signal transduction
 CC pathway, the presence of an antigen leads to the activation of an
 CC antigen-specific helper T-cell, through the action of the pathway.
 CC Undesired immune responses in a human subject (such as responses against
 CC "self-antigens") cause autoimmune diseases in humans. The peptides and
 CC peptidomimetics can therefore be used for suppressing a human CD4 T-cell
 CC mediated response. They can be used to treat autoimmune diseases such as
 CC rheumatoid arthritis, multiple sclerosis and systemic lupus
 CC erythematosus. They can also be used to prevent graft rejection.
 XX
 SQ Sequence 7 AA;

Query Match 2.4%; Score 6; DB 18; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSQ 221
 |||||
 DB 1 LSDSQ 6

RESULT 12
 AAW87684
 ID AAW87684 standard; peptide; 7 AA.
 XX
 AC AAW87684;
 XX
 DE 04-MAR-1999 (first entry)
 XX
 DE Peptide derived from D4 domain FG ridge of human CD4.
 XX
 KW human CD4; cyclic peptide; suppression; allograft rejection;
 KW CD4-dependent T-cell-mediated immune response; multiple sclerosis.
 XX
 CS Synthetic.
 OS Homo sapiens.

XX US5846933-A.
 PN 08-DEC-1998.
 PD
 XX 28-JUN-1996; 96US-0672610.
 XX 28-JUN-1996; 96US-0672610.
 XX
 XX (HUANG/) HUANG Z.
 XX (KORN/) KORNOLD R.
 XX Huang Z, Kornold R;
 XX WPI: 1999-05913/05.

XX Inhibition of T-cell immune response in vivo - by administering
 PT cyclic CD4-derived peptide(s)
 PT
 PS Disclosure: Columns 9-10; 27pp; English.
 XX

XX The present peptide is derived from human CD4. It is the core peptide of
 CC a cyclic peptide that is used in the method of the invention, which is
 CC used for suppressing a human CD4-dependent T-cell-mediated immune
 CC response. The method comprises administering a peptide of formula:
 CC R-X-Z-(CORE PEPTIDE)-Z'-X'-R' or
 CC where R-H or acetyl; R'-OH or NH2; A1 - an L-amino acid selected
 CC from N, Q and D; A2 - an L-amino acid selected from S, T, D and N;
 CC A3 - an L-amino acid selected from N, D, E and Q; A4 - an L-amino acid
 CC selected from Q, N, E and M; A5 - an L-amino acid selected from Z, L, Y
 CC and A; X - an L-amino acid or a D; or tripeptide of D- or L-amino acids
 CC selected from Y, W and F, provided that no amino acid is selected more

CC than once, or X is absent; X' - a D- or L-amino acid selected from Y,
 CC W, F, I and L or a dipeptide of D- or L-amino acids selected from L and
 CC I, or X' is absent; Z and Z' are amino acids that are linked to each
 CC other so that the peptide is cyclic; and CORE PEPTIDE is a sequence
 CC selected from AAW87645-84. The method is used especially for treating
 CC allograft rejection or multiple sclerosis.
 XX

SQ Sequence 7 AA;

Query Match 2.4%; Score 6; DB 20; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSQ 221
 |||||
 DB 1 LSDSQ 6

RESULT 13
 AAW11657
 ID AAW11657 standard; peptide; 8 AA.
 XX
 AC AAW11657;
 XX
 DE 17-OCT-1997 (first entry)
 XX
 DE Core peptide #7 for synthetic CD4 analogue.

XX CD4; therapy; macrocyclic peptidomimetic; inhibitor; MHC class II;
 KW helper T-cell; co-receptor; antigen receptor; immune response; antigen;
 KW signal transduction pathway; human; autoimmune disease; graft rejection;
 KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus.
 XX
 OS Synthetic.

XX
 XX WO9701350-A1.
 XX 16-JAN-1997.
 PD

XX 28-JUN-1996; 96WO-US11176.
 XX 20-SEP-1995; 95US-0004034.
 XX 29-JUN-1995; 95US-0000710.
 XX

(UYJE-) UNIV JEFFERSON THOMAS.

XX Huang Z, Kornold R;

XX WPI: 1997-100011/09.

XX New peptide(s) based on CD4 sequences - which suppress CD4 T-cell
 PT mediated responses, partic. for treating auto-immune disease or
 PT preventing graft rejection.
 XX

PS Claim 1; Page 34; 66pp; English.

XX AAW11651-W11694 represent core peptides that can be used in the
 CC synthetic CD4 peptides of the invention (see AAW11642-W11650). These core
 CC peptides can also be used in macrocyclic peptidomimetics of the
 CC invention. In the peptidomimetics, the core peptide is linked to a
 CC 10-member, heterocyclic ring. The peptides and peptidomimetics mimic the
 CC surface of CD4 molecules and interfere with the interaction of CD4 and
 CC MHC class II, gene products. CD4 molecules are glycoproteins expressed on
 CC the surface of helper T-cells. CD4 molecules are co-receptors for the
 CC T-cell antigen receptor, and are implicated in the signal transduction
 CC pathway for activation of helper T-cells. In the signal transduction
 CC pathway, the presence of an antigen leads to the activation of an
 CC antigen-specific helper T-cell, through the action of the pathway.
 CC Undesired immune responses in a human subject (such as responses against
 CC "self-antigens") cause autoimmune diseases in humans. The peptides and
 CC peptidomimetics can therefore be used for suppressing a human CD4 T-cell
 CC mediated response. They can be used to treat autoimmune diseases such as
 CC rheumatoid arthritis, multiple sclerosis and systemic lupus

CC erythematosis. They can also be used to prevent graft rejection.
 XX
 SQ Sequence 8 AA;

Query Match 2.4%; Score 6; DB 18; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSQ 221
 I
 Db 1 LSDSQ 6

RESULT 24
 AAW11649
 ID AAW11649 standard; peptide: 8 AA.
 XX
 AC AAW11649;
 XX
 DT 17-OCT-1997 (first entry)
 XX
 DE Synthetic CD4 analogue 34-FG.
 XX
 KW CD4: therapy; macrocyclic peptidomimetic; inhibitor; MHC class II;
 KW helper T-cell; coreceptor; antigen receptor; immune response; antigen;
 KW signal transduction pathway; human; autoimmune disease; graft rejection;
 KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus.
 XX
 OS Synthetic.
 XX
 PN WC9701350 A1.
 XX
 PD 16-JAN-1997.
 XX
 PF 28-JUN-1996; 96W0-0511176.
 XX
 PR 20-SEP-1995; 95US-0030334.
 PR 29-JUN-1995; 95US-0005710.
 XX
 PA (YJJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Huang Z, Kornagold R;
 XX
 WP1; 1997-100011/09.
 XX
 CC New peptide(s) based on CD4 sequences - which suppress CD4 T-cell
 CC mediated responses, partic. for treating auto-immune disease or
 CC preventing graft rejection.
 CC
 PS Example 6; Page 29; 56pp; English.
 XX
 CC AAW11644-W11650 represent examples of the peptides of the invention. The
 CC core peptides from these sequences (see AAW11651-W11694) can also be
 CC used in macrocyclic peptidomimetics of the invention. In the
 CC peptidomimetics, the core peptide is linked to a 10-member, heterocyclic
 CC ring. The peptides and peptidomimetics mimic the surface of CD4 molecules
 CC and interfere with the interaction of CD4 and MHC class II, gene
 CC products. CD4 molecules are glycoproteins expressed on the surface of
 CC helper T-cells. CD4 molecules are co-receptors for the T-cell antigen
 CC receptor, and are implicated in the signal transduction pathway for
 CC activation of helper T-cells. In the signal transduction pathway, the
 CC presence of an antigen leads to the activation of an antigen-specific
 CC helper T-cell, through the action of the pathway. Undesired immune
 CC responses in a human subject (such as responses against "self-antigens")
 CC cause autoimmune diseases in humans. The peptides and peptidomimetics can
 CC therefore be used for suppressing a human CD4 T-cell mediated response.
 CC They can be used to treat autoimmune diseases such as rheumatoid
 CC arthritis, multiple sclerosis and systemic lupus erythematosus. They can
 CC also be used to prevent graft rejection.
 XX
 SQ Sequence 8 AA;

Query Match 2.4%; Score 6; DB 18; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSQ 221
 I
 Db 1 LSDSQ 6

Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSQ 221
 I
 Db 1 LSDSQ 6

RESULT 15
 AAW87695
 ID AAW87695 standard; peptide: 8 AA.
 XX
 AC AAW87695;
 XX
 DT 04-MAR-1999 (first entry)
 XX
 DE Peptide derived from human CD4.
 XX
 KW Human CD4: cyclic peptide; suppression; allograft rejection;
 KW CD4-dependent T-cell-mediated immune response; multiple sclerosis.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN US5846933-A.
 XX
 PD 08-DEC-1998.
 XX
 PF 28-JUN-1996; 96US-0672610.
 XX
 PR 28-JUN-1996; 96US-0672610.
 XX
 PA (HUAN/) HUANG Z.
 PA (KORN/) KORNAGOLD R.
 XX
 PI Huang Z, Kornagold R;
 XX
 DR WP1; 1999-059113/05.
 XX
 PT Inhibition of T-cell immune response in vivo - by administering
 PT cyclic CD4-derived peptide(s)
 XX
 PS Example 6; Column 18; 27pp; English.
 XX
 CC The present peptide is derived from human CD4. The specification
 CC describes a method for suppressing a human CD4-dependent T-cell-mediated
 CC immune response. The method comprises administering a peptide of
 CC formula: R-X-Z-Al-A2-A3-A4-A5-Z'-X'-R' or
 CC R-X-Z-(CORE PEPTIDE)-Z'-X'-R'
 CC where R= H or acetyl; R'= OH or NH2; A1 = an L-amino acid selected
 CC from N, Q and D; A2 = an L-amino acid selected from N, D, E and Q; A4 = an L-amino acid
 CC selected from Q, N, E and M; A5 = an L-amino acid selected from I, L, V
 CC and A; X= an L-amino acid or a di- or tripeptide of D- or L-amino acids
 CC selected from Y, W and F; provided that no amino acid is selected more
 CC than once, or X is absent; X'= a D- or L-amino acid selected from Y,
 CC W, F, I and L, or a dipeptide of D- or L-amino acids selected from Y,
 CC L, or X' is absent; Z and Z' are amino acids that are linked to each
 CC other so that the peptide is cyclic; and CORE PEPTIDE is a sequence
 CC selected from AAW87645-84. The method is used especially for treating
 CC allograft rejection or multiple sclerosis.
 XX
 SQ Sequence 8 AA;

Query Match 2.4%; Score 6; DB 20; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSQ 221
 I
 Db 1 LSDSQ 6

```

RESULT 15
ID AAW87650 standard: peptide; 8 AA.
XX
AC AAW87650:
XX
DT 04-MAR-1999 (first entry)
XX
DE Peptide derived from E4 domain FG ridge of human CD4.
XX
KW Human CD4; cyclic peptide; suppression; allograft rejection;
KW CD4-dependent T-cell-mediated immune response; multiple sclerosis.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN US9846533-A.
XX
PD 08-DEC-1998.
XX
PF 28-JUN-1996; 960S 0672620.
XX
PR 28-JUN-1996; 960S-0672620.
XX
PA (HUANG) HUANG Z.
PA (KORN) KORNOLD R.
XX
PI Huang Z, Kornold R;
XX
WPI: 1999-059113/05.
XX
PI Inhibition of T cell immune response in vivo - by administering
PI cyclic CD4-derived peptide(s)
XX
PS Disclosure: Column 8; 27pp: English.
XX
CC The present peptide is derived from human CD4. It is the core peptide of
CC a cyclic peptide that is used in the method of the invention, which is
CC used for suppressing a human CD4-dependent T-cell-mediated immune
CC response. The method comprises administering a peptide of formula:
CC R-X-Z-A1-A2-A3-A4-A5-Z'-X'-R' or
CC R-X-Z-(CORE PEPTIDE)-Z'-X'-R'
CC where R= H or acetyl; R'= OH or NH2; A1 = an L-amino acid selected
CC from N, Q and D; A2 = an L-amino acid selected from S, T, D and N;
CC A3 = an L-amino acid selected from N, D, E and Q; A4 = an L-amino acid
CC selected from Q, N, E and M; A5 = an L-amino acid selected from I, L, V
CC and A; X= an L-amino acid or a di- or tripeptide of D- or L-amino acids
CC selected from Y, W and F; provided that no amino acid is selected more
CC than once, or X is absent; X'= a D- or L-amino acid selected from Y,
CC W, F, I and L or a dipeptide of D- or L-amino acids selected from L and
CC I, or X' is absent; Z and Z' are amino acids that are linked to each
CC other so that the peptide is cyclic; and CORE PEPTIDE is a sequence
CC selected from AAW87645-84. The method is used especially for treating
CC allograft rejection or multiple sclerosis.
XX
SQ Sequence 8 AA:
Query Match 2.4%; Score 6; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 LSDSQ 221
DB 1 LSDSQ 6
|||||

RESULT 17
ID AAR83816 standard: peptide; 9 AA.
XX
AC AAR83816:
XX
DT 14-FEB-1996 (first entry)
XX
DE SH3-binding proline-rich p47(phox) analogue #9.
XX
KW p47(phox); p67(phox); NADPH oxidase complex; proline-rich region; PTPase;
KW SH3 domain; inhibition; dynamin; cytochrome b245; reperfusion injury;
KW septic shock; arthritis; asthma; vinculin; inflammatory bowel disease;
KW adult respiratory distress syndrome; ischaemic heart disease.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN GB2285047-A.
XX
PD 28-JUN-1995.
XX
PF 07-DEC-1994; 94GB-0024674.
XX
PR 07-JAN-1994; 94GB-0000248.
PR 21-DEC-1993; 93GB-0026083.
XX
PA (LUDWIG) LUDWIG INST CANCER RES.
PA (YAMA) YAMANOUCHI UK LTD.
XX
PI Finan PM, Gout IT, Keillie S, Shimizu Y, Waterfield MD;
XX
WPI: 1995-217703/29.
XX
PI New polypeptide(s) with antiinflammatory action - inhibit NADPH
PI oxidase system.
XX
PS Claim 2; Page 8; 17pp: English.
XX
CC The protein p47(phox) (AAR83825) interacts with protein p67(phox) in the
CC NADPH oxidase complex via a C-terminal proline-rich region of p47(phox)
CC binding to at least one C-terminal SH3 domain of p67(phox). The
CC peptides AAR83814-24 were derived from the proline-rich region of
CC p47(phox) and show inhibitory activity towards the binding of p47(phox)
CC to p67(phox) as compared to the proline-rich region of a range of other
CC proteins (AAR83826-30). The inhibitory peptides can be used for the
CC treatment of chronic or acute inflammatory diseases e.g. septic shock,
CC arthritis, asthma, adult respiratory distress syndrome, ischaemic heart
CC disease, reperfusion injury or inflammatory bowel disease.
XX
SQ Sequence 9 AA:
Query Match 2.4%; Score 6; DB 16; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 PAVPPH 57
DB 2 PAVPPH 7
|||||

RESULT 18
ID AAR83822 standard: peptide; 9 AA.
XX
AC AAR83822:
XX
DT 14-FEB-1996 (first entry)
XX
DE SH3-binding proline-rich p47(phox) analogue #9.
XX
KW p47(phox); p67(phox); NADPH oxidase complex; proline-rich region; PTPase;
KW SH3 domain; inhibition; dynamin; cytochrome b245; reperfusion injury;
KW septic shock; arthritis; asthma; vinculin; inflammatory bowel disease;
KW adult respiratory distress syndrome; ischaemic heart disease.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN GB2285047-A.
XX
PD 28-JUN-1995.

```

XX PF 07-DEC-1994: 94GB-0024674.
 XX XX
 XX PR 07-JAN-1994: 94GB-0000248.
 XX PR 21-DEC-1993: 93GB-0026683.
 XX XX
 XX PA (UDW-) IUDWIG INST CANCER RES.
 XX PA (YAMA) YAMANOUCHI UK LTD.
 XX XX
 XX PT Firan PM, Gout H, Kellie S, Shimizu Y, Waterfield MD;
 XX WPI: 1995-2.7753/29.
 XX DR WPI: 1995-2.7753/29.
 XX XX
 XX PF New polypeptide(s) with antiinflammatory action - inhibit NADPH
 XX PF oxidase system.
 XX XX
 XX PS Claim 2; Page 8; 17pp; English.
 XX XX
 XX CC The protein p47(phox) (AA83825) interacts with protein p67(phox) in the
 CC NADPH oxidase complex via a C-terminal proline-rich region of p47(phox)
 CC binding to at least one C-terminal SH3 domain of p67(phox). The
 CC peptides AA83614-24 were derived from the proline-rich region of
 CC p47(phox) and show inhibitory activity towards the binding of p47(phox)
 CC to p67(phox) as compared to the proline-rich region of a range of other
 CC proteins (AA83626-30). The inhibitory peptides can be used for the
 CC treatment of chronic or acute inflammatory diseases e.g. septic shock,
 CC arthritis, asthma, adult respiratory distress syndrome, ischaemic heart
 CC disease, reperfusion injury or inflammatory bowel disease.
 XX XX
 XX SQ Sequence 9 AA;
 Query Match 2.4%; Score 6; DB 16; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e-05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 YFAVPP 56
 DB 1 YFAVPP 6
 II.III
 AA098604
 ID AA098604 standard; Peptide: 11 AA.
 XX AC
 XX AC AA098804;
 XX DT 24-JAN 2002 (first entry)
 XX DE Human peptide #2079 encoded by a SNP oligonucleotide.
 XX XX
 KW Immunopressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease.
 XX XX
 OS Homo sapiens.
 XX XX
 XX PN W0200147944 A2.
 XX PD 35-JUL 2001.
 XX XX
 XX PE 28-DEC-2000; 2000WJ-US35498.
 XX PR 28-DEC-1999; 990US-0173419.
 XX PR 27-DEC-2000; 2000US-0173419.
 XX XX
 XX PA (CURA-) CURAGEN CORP.
 XX XX
 XX PT Shamkels RA, Leach M;

XX DR WPI: 2001-465210/50.
 XX XX
 XX PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 XX PT oncogenes and histones, useful for diagnosing and treating, e.g.
 XX PT cancer, autoimmune diseases and infections.
 XX PS Disclosure: Page 4124; 4143pp; English.
 XX XX
 XX CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, cytochromes, kinesins, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.
 XX XX
 XX SQ Sequence 11 AA;
 Query Match 2.4%; Score 6; DB 22; Length 11;
 Best Local Similarity 100.0%; Pred. No. 1.1e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 SUPEGI 12
 DB 1 SUPEGI 6
 II.III
 AA03416
 ID AA03416 standard; peptide: 12 AA.
 XX AC
 XX AC AA03416;
 XX DT 09-OCT-1996 (first entry)
 XX DE CDR3-like constrained peptide.
 XX XX
 KW constrained; CD4; complementarity determining region; CDR;
 KW HIV; gp120; lymphoma; inflammation; cyclic.
 XX OS Synthetic.
 XX XX
 XX FH Location/Qualifiers
 XX FT Key Region 1
 FT Disulfide-bond 2..11 /note- "exocyclic aromatic region"
 FT /note- "the peptide is constrained into a loop by
 FT this bond"
 FT Region 3..10
 FT /label- human_CD4_(349-356)
 FT Region 12 /note- "this is a biologically active loop region"
 FT /note- "exocyclic aromatic region"
 XX PN W09534312-A1.
 XX XX
 XX PD 21-DEC-1995.
 XX XX
 XX PF 05-JUN-1995; 95WQ-US07157.
 XX XX
 XX PR 10-JUN-1994; 94US-0257783.
 XX XX
 XX PA (UYPE-) UNIV PENNSYLVANIA.

```

XX  Greene ML;
XX
XX  WPI: 1996-043412/05.
XX
XX  Novel constrained peptide having biologically active loop region -
XX  contains aromatic amino acids at exocyclic positions to improve
XX  interaction of loop region with other molecules, e.g. receptors or
XX  ligands
XX
XX  Disclosure; Page 19; 38pp; English.
XX
XX  New aromatically modified constrained peptides are provided which (1)
XX  consist of 7-30 amino acids, (2) have a central cyclic portion
XX  composed of a biologically active sequence of 3-18 amino acids, and (3)
XX  have two terminal exocyclic portions each of 1-6 amino acids in length
XX  and each containing at least one aromatic amino acid, specifically
XX  tyrosine or phenylalanine. Preferably the biologically active sequence
XX  consists of an antibody loop (CDR) or a CDR-like loop from a receptor or
XX  other member of the immunoglobulin superfamily, and the cyclisation is
XX  due to a disulphide bond. The loop region is able to affect processes
XX  involved in e.g. lymphoma cell growth, inflammation, binding of HIV
XX  gp120 to CD4, etc. The peptides show enhanced binding to a target
XX  region compared with the corresponding linear peptide, and the aromatic
XX  amino acids in the exocyclic portions of the molecule provide increased
XX  interactions. The present sequence represents one of 8 exemplified such
XX  peptides.
XX
XX  Query Match: 2.4%; Score 6; DB 17; Length 12;
XX  Best local Similarity 100.0%; Pred. No. 1.1e+02;
XX  Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  QY 216 LSDSQ 221
XX  IIIII
XX  3 LSDSQ 8
XX
XX  RESULT 21
XX  AAW26034
XX  ID AAW26034 standard; peptide; 12 AA.
XX
XX  AC AAW26034;
XX
XX  DT 28-OCT-1997 (first entry)
XX
XX  DE CDK8 derivative of tenezin; peptide fragment TED.
XX
XX  KW Tenezin; antibiotic; antifungal; peptide; Tenebrio molitor; chemotherapy;
XX  systemic infection; pathogen.
XX
XX  OS Synthetic.
XX
XX  FH Key Location/Qualifiers
XX  FT Modified-site 12 /note= "amidated"
XX
XX  PN W09702286-A1.
XX
XX  PU 23-JAN-1997.
XX
XX  FF 11-MAR-1996; 96WO 060034.
XX
XX  PR 29-JAN-1996; 96KR 0001911.
XX  PR 06-JUL-1995; 95KR 0015694.
XX  PR 29-JAN-1996; 96KR 0001909.
XX  PR 29-JAN-1996; 96KR 0001910.
XX
XX  PA (MOGA-) MOSAM BIOTECHNOLOGY RES CNS1.
XX
XX  PI Che H, Chung K, Hong S, Lee H, Lee K, Moon H;
XX  Oh J, Yoon J;

```

```

XX  WPI: 1997-108913/10.
XX
XX  Acid- or amide-form peptide(s) with antibacterial and antifungal
XX  activity - used for chemotherapy of local and systemic infections
XX  caused by pathogenic bacteria
XX
XX  Example 2; Page 14; 30pp; English.
XX
XX  AAW26002-W26077 represent derivatives of the antibiotic fragments of
XX  tenezin (see AAW26000) shown in AAW01830-W10835. AAW01830-W10835 are
XX  amidated derivatives of antibiotic fragments of the wild type tenezin
XX  sequence. These sequences are used as the antibacterial and antifungal
XX  peptides of the invention. Tenezin is an antibacterial peptide isolated
XX  from Tenebrio molitor larvae. Tenezin does have some drawbacks which
XX  prevent it from practical use. Tenezin has a narrow spectrum of target
XX  cells, and due to its large molecular size may provoke antigen-antibody
XX  reactions in vivo, and is also unstable. The peptides can be used for the
XX  development of antibacterial and antifungal agents for the chemotherapy
XX  of local and systemic infections caused by pathogenic bacteria and/or
XX  fungi and can be formulated into potent antibacterial and/or fungal
XX  agents. The peptides have superior antibacterial and/or antifungal
XX  activity, while causing no cytotoxicity. They do not give rise to lysing
XX  of red blood cells. These peptides also have improved stability over the
XX  wild type tenezin.
XX
XX  Query Match: 2.4%; Score 6; DB 18; Length 12;
XX  Best local Similarity 100.0%; Pred. No. 1.1e+02;
XX  Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  QY 75 NGKVCV 80
XX  IIIII
XX  5 NGKVCV 10
XX
XX  RESULT 22
XX  AAW15302
XX  ID AAW15302 standard; peptide; 14 AA.
XX
XX  AC AAW15302;
XX
XX  DT 01-JUL-1997 (first entry)
XX
XX  DE Placental lactogen whose release is decreased in hypoxia.
XX
XX  KW Release; expression; secretion; mammary; foetus; trophoblast; cell;
XX  chorionic villus; decrease; hypoxia; marker; indicator; abnormal;
XX  maternal; placental; interface; function; abortion; screening;
XX  intrauterine; growth; retardation; gestation; disease;
XX  molar; pregnancy; choriocarcinoma; tumour; ectopic; lactogen;
XX  proteinuria; hypertension; preeclampsia; induction; mitigation.
XX
XX  OS Homo sapiens.
XX
XX  PN W09633214-A2.
XX
XX  PU 24-OCT-1996.
XX
XX  PF 18-APR-1996; 96WO-US05441.
XX
XX  PR 18-APR-1995; 95US-0423409.
XX
XX  PA (REGC ) UNIV CALIFORNIA.
XX
XX  PI Burlingame AL, Clauser KR, Fisher SJ, Foulk R, Genbacev O;
XX  WPI: 1996-497265/49.
XX
XX  PR Detection of abnormal placental function or metastases - by
XX  detecting proteins with altered expression or trophoblasts or
XX  chorionic villi under hypoxic conditions

```

XX
PS Claim 1: Page 40; 57pp; English.
XX
CC The level of release of the present peptide by a mammalian foetal
CC trophoblast cell or a chorionic villus is decreased when the cell
CC or villus is grown under hypoxic conditions, characterised by a
CC partial pressure of oxygen (pO₂) of 14 mm Hg. The peptide can be
CC used as a marker for the presence of hypoxic conditions indicative
CC of an abnormal maternal-placental interface, and consequent
CC abnormal placental function in, e.g. threatened abortion,
CC intra-uterine growth retardation, gestational trophoblast diseases
CC including molar pregnancy, choriocarcinoma, placental site tumours,
CC ectopic pregnancy, proteinuria, pregnancy induced hypertension and
CC preeclampsia. It can also be used to screen for inducers or
CC mitigators of abnormal maternal-placental interface.
XX
SQ Sequence 14 AA;

Query Match 2.4%; Score 6; DB 17; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 CESLE 186
DE 6 IESWLE 11
IIII

RESULT 23
AB054:4
ID AB054:4 standard; Peptide; 14 AA.
XX AC AB054:4;
XX
XX 11-APR-2002 (first entry);
XX
XX Membrane occupation and recognition nexus motif peptide SEQ ID NO:23.
XX
XX Rabbit; Juncetophilin; JP 1: physiological transfer structure;
KW dihydropyridine receptor; lyansine receptor; MORN motif; MORN;
KW membrane occupation and recognition nexus.
XX
XX Arabidopsis thaliana.
OS
XX JP2503321172-A.
PN
XX 20-NOV-2001.
PD
XX 10-MAY-2000; 2003JP-0137555.
PF
XX 10-MAY-2000; 2003JP-0137555.
PR
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
PA
XX WPI; 2002-144132/19.
DR
XX
XX New gene encoding protein molecule Juncetophilin -
PT
XX
XX Claim 29; Page 22; 20pp; Japanese.
PS
XX
XX The present invention describes rabbit juncetophilin (JP-1). The present
CC invention also describes: (1) a fused protein in which a protein having
CC juncetophilin activity is combined with a marker protein and/or a peptide
CC tag; (2) an antibody, preferably a monoclonal antibody, specifically
CC combining with a protein having juncetophilin activity; (3) a hybridoma
CC producing the above monoclonal antibody; (4) a host cell containing an
CC expression system which can express a protein having juncetophilin
CC activity; (5) a nonhuman animal in which the gene function encoding a
CC protein having juncetophilin activity is deleted or excessively expressed
CC on chromosomes; and (6) preparation of cells to which juncetophilin
CC activity is given by introducing a juncetophilin encoding polynucleotide
CC in to cells in which the gene function encoding a protein having
CC juncetophilin activity is deleted on chromosomes. The juncetophilin
CC protein, gene or DNA can be used for obtaining information on

XX
PS physiological transfer structure between a dihydropyridine receptor and
XX a lyansine receptor. The present sequence represents a specifically
CC claimed membrane occupation and recognition nexus (MORN) motif peptide
CC from the present invention.
XX
SQ Sequence 14 AA;

Query Match 2.4%; Score 6; DB 23; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 RGLQEG 132
DE 7 RGLQEG 12
IIII

RESULT 24
AAW99007
ID AAW99007 standard; peptide; 16 AA.
XX AC AAW99007;
XX
XX 11-MAY-1999 (first entry)
DT
XX I domain peptide P12.
DE
XX Cyclic integrin binding peptide; integrin alpha-2I domain; inhibitor;
KW collagen 2; collagen IV; laminin-1; cell migration; cancer;
KW cardiovascular disease; periodontal disease.
XX
XX Synthetic.
OS
XX WO9902551-A1.
PN
XX 21-JAN-1999.
PD
XX 09-JUL-1998; 98WO-F100579.
PF
XX 11-JUL-1997; 97US-0893526.
PR
XX (HEIN/) HEINO J.
PA (IVAS/) IVASKA J.
PA (KAPP/) KAEPLYAE J.
XX
XX Heino J, Ivaska J, Kaepylae J;
WPI; 1999-120775/10.
DR
XX Cyclic integrin binding peptides - used to inhibit
PT integrin-dependent cell migration
PT
XX
XX Example 3; Page 25; 59pp; English.
PS
XX
XX The present sequence invention describes cyclic integrin binding peptides
CC comprising the amino acid sequence RKK, preferably RKKH. Also described
CC in the present invention are: (1) a cyclic peptide comprising the amino
CC acid sequence XIRKKHX2Xn where X is any amino acid and n=1-4; (2) a
CC cyclic integrin binding peptide comprising the amino acid sequence
CC CTRKKHNC or CTRKKHNAQC; (3) a pharmaceutical composition comprising
CC one of the above integrin binding peptides; and (4) a binding assay for
CC identifying integrin binding agents, comprising: (a) biotinylating the
CC integrin binding agent to be assayed; (b) reacting the biotinylated
CC agent with an immobilised recombinant alpha 2I domain or domain-derived
CC peptides in conditions suitable for binding; (c) washing the solid
CC support carrying the bound agent; (d) adding a labelled biotin-binding
CC agent; and (e) detecting any bound integrin binding agent. The integrin
CC binding peptides can be used for inhibiting integrins. They can also be
CC used for inhibiting integrin-dependent cell migration, such as
CC associated with cancer, cardiovascular disease or a periodontitis
CC condition. They can also be used for inhibiting the adhesion of
CC platelets to collagen or collagen induced platelet aggregation in a
CC patient. The present sequence represents a peptide from the present
CC invention.

XX Sequence 16 AA;
 Query Match 2.4%; Score 6; DB 20; Length 16;
 Best local similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 ALLEKA 194
 11111
 DB 5 ALLEKA 10

RESULTS 25
 AAB70014
 ID AAB70014 standard; Peptide: 16 AA.
 AC AAB70014:
 XX
 XX
 XX 08-MAY-2001 (first entry)
 XX
 XX Mammalian CD4 IL-16 antagonist longer peptide, SEQ ID NO: 12.
 XX Human; mouse; CD4; Interleukin-16; IL-16; IL-16 antagonist;
 KW antiasthmatic; antiarthritic; antiinflammatory; antithyroid;
 KW neuroprotective; anti-allergic; ophthalmological; dermatological;
 KW immunosuppressive; inflammatory disease; asthma; arthritis;
 KW inflammatory bowel disease; Grave's ophthalmopathy; multiple sclerosis;
 KW systemic lupus erythematosus; SLE; atopic rhinitis; atopic dermatitis;
 KW bullous pemphigoid.
 XX
 XX Homo sapiens.
 OS Mus sp.
 XX
 XX W0200110891-A2.
 XX 15-FEB-2001.
 XX 04-AUG-2000; 2000WO US21224.
 XX 05-AUG-1999; 9905-0368632.
 XX (RFSF) RESEARCH CORP TECHNOLOGIES INC.
 XX
 XX Crichtonshank WW, Kornfeld H, Center DM;
 XX WPI: 2001 202756/20.
 XX Novel interleukin-16 antagonist peptide useful for treating
 XX inflammatory diseases such as asthma, arthritis, Grave's ophthalmopathy,
 XX multiple sclerosis, lupus, atopic rhinitis, atopic dermatitis or
 XX bullous pemphigoid -
 XX
 XX Claim 4; Page 57; 75pp; English.
 XX
 XX The present sequence is provided in a specification relating to novel
 XX interleukin-16 (IL-16) antagonist peptides. The peptides are useful for
 XX treating an IL-16-mediated disorder, especially inflammatory diseases
 XX such as asthma, arthritis, inflammatory bowel disease, Grave's
 XX ophthalmopathy, multiple sclerosis, systemic lupus erythematosus
 XX (SLE), atopic rhinitis, atopic dermatitis or bullous pemphigoid.
 XX The peptides comprise specific portions of the native human CD4
 XX receptor and variants of the receptor and are therefore non-immunogenic
 XX when administered to human. They act by blocking the interaction of
 XX IL-16 with an IL-16 or CD4 receptor and may be used in conjunction with
 XX an anti-inflammatory agent, such as an anti-CD4 antibody, an anti-tumour
 XX necrosis factor (TNF)-alpha antibody, non-steroidal antiinflammatory
 XX drugs, steroids or cyclosporin-A.
 XX
 XX Sequence 16 AA;
 Query Match 2.4%; Score 6; DB 22; Length 16;
 Best local similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSG 221
 11111
 DB 7 LSDSG 12

RESULTS 26
 ABG66616
 ID ABG66616 standard; Peptide: 19 AA.
 XX
 XX AC ABG66616;
 XX
 XX 30-AUG-2002 (first entry)
 XX
 XX IqE Fcepsilon RI binding peptide IGE134 721-13.
 XX
 XX IqE receptor; immunoglobulin; FcepsilonRI; antagonist; phage display;
 KW protein co-ordinate data; IqE-mediated disease; allergic rhinitis;
 KW asthma; allergic asthma; atopic dermatitis; urticaria-angioedema;
 KW parasitic infection; IqE myeloma; immune-related disorder;
 KW inflammatory disorder; diabetes mellitus; reperfusion injury; stroke;
 KW IqE-mediated gastrointestinal inflammatory disease; burn;
 KW immune rejection of graft; myocardial infarction; atherosclerosis;
 KW acute lung injury; haemorrhagic shock; septic shock;
 KW acute tubular necrosis; endometriosis; degenerative joint disease;
 KW pancreatitis.
 XX
 XX Synthetic.
 XX
 XX W0200226781-A2.
 XX 04-APR-2002.
 XX 26-SEP-2001; 2001WO-US30289.
 XX 26-SEP-2000; 2000US-235353P.
 XX 23-MAR-2000; 2000US-278540P.
 XX (GETH) GENENTECH INC.
 XX
 XX Lowman HB, Reynolds ME, Nakamura GR, Starovasnik MA;
 XX WPI: 2002-444016/47.
 XX A peptide useful for treating a IqE-mediated disease or disorder in a
 XX host e.g. allergic rhinitis, asthma, which competes with immunoglobulin
 XX E for binding to high affinity IqE receptor in an in vitro assay -
 XX
 XX Example 9; Fig 5; 328pp; English.
 XX
 XX The invention relates to a peptide which competes with immunoglobulin
 XX (Ig) F134 comprising a sequence (SI), for binding the high affinity IqE
 XX receptor (FcepsilonRI) in an in vitro assay and having a formula given in
 XX the specification. Also included are a fusion protein comprising the
 XX peptide, a pharmaceutical composition (C) comprising the peptide,
 XX designing a compound that mimics the three-dimensional surface
 XX structure of the peptide, a compound with a solvent accessible surface
 XX that mimics the solvent accessible surface defined by the side chains of
 XX residues (R) Pro4, Phe6, Pro16, Cys7, Cys15 and Cys19 of IqE134,
 XX a peptide with structural coordinates as given in the specification,
 XX selecting a peptide mimetic which binds to FcepsilonRI and blocks
 XX binding of IqE and a peptide mimetic which mimics the coordinates of
 XX IqE134 residues (R). (C) is useful for inhibiting the binding of IqE to
 XX high affinity IqE receptor (FcepsilonRI). Peptides of the formula given
 XX in the specification are useful for inhibiting the binding of an IqE to
 XX high affinity IqE receptor. The peptide is useful for selecting a
 XX molecule which blocks the interaction of IqE with high affinity IqE
 XX receptor. The peptide is also useful for inhibiting the activation of
 XX high affinity IqE receptor. The peptide is useful for treating an IqE-
 XX mediated disease or disorder in a host. (C) is useful in research,
 XX diagnostic, therapeutic and prophylactic methods. The peptide is also
 XX useful for inhibiting IqE-mediated or associated processes such as IqE-
 XX dependent activation and degranulation of mast cells and basophils, as

CC well as consequent release of inflammatory mediators such as histamine.
 CC (C) is useful for treating allergic rhinitis, asthma (e.g. allergic
 CC asthma), atopic dermatitis, urticaria-angioedema, parasitic infection,
 CC IgE myeloma, immune-related disorders, inflammatory disorders, diabetes
 CC mellitus, IgE-mediated gastrointestinal inflammatory disease, immune
 CC rejection of grafts, reperfusion injury, stroke, myocardial infarction,
 CC atherosclerosis, acute lung injury, haemorrhagic shock, burn, septic
 CC shock, acute tubular necrosis, endometriosis, degenerative joint disease
 CC and pancreatitis. The present sequence is a peptide of the invention
 CC expressed from a phage display library.
 XX
 XX Sequence 19 AA;

Query Match 2.4%; Score 6; DB 23; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.7e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 PHECYL 6;

Db 4 PHECYL 9

RESULT 27

ABG66605
 ID ABG66605 standard; Peptide: 20 AA.

XX
 AC ABG66605;

DT 30-AUG-2002 (first entry)

DE IgE Fcεpsilon RI binding peptide IGE134 720-2.

XX IgE receptor; immunoglobulin; FcεpsilonRI; antagonist; phage display;
 KW protein co-ordinate data; IgE-mediated disease; allergic rhinitis;
 KW asthma; allergic asthma; atopic dermatitis; urticaria-angioedema;
 KW parasitic infection; IgE myeloma; immune-related disorder;
 KW inflammatory disorder; diabetes mellitus; reperfusion injury; stroke;
 KW IgE-mediated gastrointestinal inflammatory disease; burn;
 KW immune rejection of graft; myocardial infarction; atherosclerosis;
 KW acute lung injury; haemorrhagic shock; septic shock;
 KW acute tubular necrosis; endometriosis; degenerative joint disease;
 KW pancreatitis.

XX Synthetic.

XX WO200226781-A2.

XX 04-APR-2002.

XX 26-SEP-2001; 2001WO-US30289.

XX 26-SEP-2000; 2000US-235353P.

XX 23-MAR-2001; 2001US-278540P.

XX (GETH) GENENTECH INC.

XX Lowman HB, Reynolds ME, Nakamura GR, Starovasnik MA;

XX WPI: 2002-444016/47.

XX A peptide useful for treating a IgE-mediated disease or disorder in a
 XX host e.g. allergic rhinitis, asthma, which competes with immunoglobulin
 XX E for binding to high affinity IgE receptor in an in vitro assay -

XX Example 9; Fig 5; 328pp; English.

XX The invention relates to a peptide which competes with immunoglobulin
 XX (Ig) E 134 comprising a sequence (SJ), for binding the high affinity IgE
 XX receptor (FcεpsilonRI) in an in vitro assay and having a formula given in
 XX the specification. Also included are a fusion protein comprising the
 XX peptide, a pharmaceutical composition (C) comprising the peptide,
 XX designing a compound that mimics the three-dimensional surface
 XX structure of the peptide, a compound with a solvent accessible surface

CC that mimics the solvent accessible surface defined by the side chains of
 CC residues (R) Pro4, Phe6, Pro16, Cys3, Cys7, Cys15 and Cys19 of IGE134,
 CC a peptide with structural coordinates as given in the specification,
 CC selecting a peptide mimetic which binds to FcεpsilonRI and blocks
 CC binding of IgE and a peptide mimetic which mimics the coordinates of
 CC IGE134 residues (R). (C) is useful for inhibiting the binding of IgE to
 CC high affinity IgE receptor (FcεpsilonRI). Peptides of the formula given
 CC in the specification are useful for inhibiting the binding of an IgE to
 CC high affinity IgE receptor. The peptide is useful for selecting a
 CC molecule which blocks the interaction of IgE with high affinity IgE
 CC receptor. The peptide is also useful for inhibiting the activation of
 CC high affinity IgE receptor. The peptide is useful for treating an IgE-
 CC mediated disease or disorder in a host. (C) is useful in research,
 CC diagnostic, therapeutic and prophylactic methods. The peptide is also
 CC useful for inhibiting IgE-mediated or associated processes such as IgE-
 CC dependent activation and degranulation of mast cells and basophils, as
 CC well as consequent release of inflammatory mediators such as histamine.
 CC (C) is useful for treating allergic rhinitis, asthma (e.g. allergic
 CC asthma), atopic dermatitis, urticaria-angioedema, parasitic infection,
 CC IgE myeloma, immune-related disorders, inflammatory disorders, diabetes
 CC mellitus, IgE-mediated gastrointestinal inflammatory disease, immune
 CC rejection of grafts, reperfusion injury, stroke, myocardial infarction,
 CC atherosclerosis, acute lung injury, haemorrhagic shock, burn, septic
 CC shock, acute tubular necrosis, endometriosis, degenerative joint disease
 CC and pancreatitis. The present sequence is a peptide of the invention
 CC expressed from a phage display library.
 XX

SQ Sequence 20 AA;

Query Match 2.4%; Score 6; DB 23; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.5e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 PHECYL 61

Db 4 PHECYL 9

RESULT 28

ABG66606

ID ABG66606 standard; Peptide: 20 AA.

XX
 AC ABG66606;

XX 30-AUG-2002 (first entry)

XX IgE Fcεpsilon RI binding peptide IGE134 720-7.

XX IgE receptor; immunoglobulin; FcεpsilonRI; antagonist; phage display;
 KW protein co-ordinate data; IgE-mediated disease; allergic rhinitis;
 KW asthma; allergic asthma; atopic dermatitis; urticaria-angioedema;
 KW parasitic infection; IgE myeloma; immune-related disorder;
 KW inflammatory disorder; diabetes mellitus; reperfusion injury; stroke;
 KW IgE-mediated gastrointestinal inflammatory disease; burn;
 KW immune rejection of graft; myocardial infarction; atherosclerosis;
 KW acute lung injury; haemorrhagic shock; septic shock;
 KW acute tubular necrosis; endometriosis; degenerative joint disease;
 KW pancreatitis.

XX Synthetic.

XX WO200226781-A2.

XX 04-APR-2002.

XX 26-SEP-2001; 2001WO-US30289.

XX 26-SEP-2000; 2000US-235353P.

XX 23-MAR-2001; 2001US-278540P.

XX (GETH) GENENTECH INC.

XX Lowman HB, Reynolds ME, Nakamura GR, Starovasnik MA;

XX WPI: 2002-444016/47.

XX A peptide useful for treating a IgE-mediated disease or disorder in a

XX host e.g. allergic rhinitis, asthma, which competes with immunoglobulin

XX E for binding to high affinity IgE receptor in an in vitro assay

XX

XX Example 3; Fig 5; 328pp; English.

XX The invention relates to a peptide which competes with immunoglobulin

XX (IgE) comprising a sequence (S1), for binding the high affinity IgE

XX receptor (FcεR1) in an in vitro assay and having a formula given in

XX the specification. Also included are a fusion protein comprising the

XX peptide, a pharmaceutical composition (C) comprising the peptide,

XX designing a compound that mimics the three-dimensional surface

XX structure of the peptide, a compound with a solvent accessible surface

XX that mimics the solvent accessible surface defined by the side chains of

XX residues (R) Pro4, Phe6, Cys3, Cys7, Cys15 and Cys19 of IgE134,

XX a peptide with structural coordinates as given in the specification,

XX selecting a peptide mimetic which binds to FcεR1 and blocks

XX binding of IgE and a peptide mimetic which mimics the coordinates of

XX IgE134 residues (R). (C) is useful for inhibiting the binding of IgE to

XX high affinity IgE receptor (FcεR1). Peptides of the formula given

XX in the specification are useful for inhibiting the binding of an IgE to

XX high affinity IgE receptor. The peptide is useful for selecting a

XX molecule which blocks the interaction of IgE with high affinity IgE

XX receptor. The peptide is also useful for inhibiting the activation of

XX high affinity IgE receptor. The peptide is useful for treating an IgE-

XX mediated disease or disorder in a host. (C) is useful in research,

XX diagnostic, therapeutic and prophylactic methods. The peptide is also

XX useful for inhibiting IgE-mediated or associated processes such as IgE-

XX dependent activation and degranulation of mast cells and basophils, as

XX well as consequent release of inflammatory mediators such as histamine.

XX (C) is useful for treating allergic rhinitis, asthma (e.g. allergic

XX asthma), atopic dermatitis, urticaria-angioedema, parasitic infection,

XX IgE myeloma, immune-related disorders, inflammatory disorders, diabetes

XX mellitus, IgE-mediated gastrointestinal inflammatory disease, immune

XX rejection of grafts, reperfusion injury, stroke, myocardial infarction,

XX atherosclerosis, acute lung injury, haemorrhagic shock, burn, septic

XX shock, acute tubular necrosis, endometriosis, degenerative joint disease

XX and pancreatitis. The present sequence is a peptide of the invention

XX expressed from a phage display library.

XX

XX Sequence 20 AA:

Query Match 2.4%; Score 6; DB 23; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 PRCV1 61

DB III II

4 PRCV1 9

RESULT 29

AAU93306

AAU93306 standard; Peptide: 20 AA.

AAU93306;

02-JUL-2002 (first entry)

Granulocyte-colony stimulating factor receptor binding peptide #112.

G-CSF; granulocyte-colony stimulating factor receptor; cytokine;

haematopoietic growth factor; neutrophil proliferation; AIDS;

neutrophil differentiation; acquired immunodeficiency syndrome;

chemotherapy-induced neutropenia; community acquired pneumonia;

depressed neutrophil count; immunostimulant.

Synthetic.

W020207676-A2.

31-JAN-2002.

20-JUL-2001; 2001WO-US23046.

20-JUL-2000; 2000US-0620051.

(GLAXO) GLAXO GROUP LTD.

Cwiria SE, Balu P, Duffin DZ, Piplani S, McHown-Merrill H;

Schatz PJ;

WPI: 2002-329382/36.

Novel compounds, useful for treating depressed neutrophil count,

comprise peptide chains of approximately 6 to 40 amino acids in length

that bind to granulocyte-colony stimulating factor receptor.

Disclosure; Fig 1; 90pp; English.

The invention relates to compounds comprising a peptide chain

approximately 6 to 40 amino acids in length that binds to granulocyte-

colony stimulating factor receptor (G-CSFR). The compounds contain

specific sequences of the generic peptides appearing as AAU79402-AAU79406

and the generic sequences XV_1XV_2XV_3XV_4XV_5XV_6XV_7XV_8 (where

XV_1 - E, C, Q, V or Y; XV_2 - E, A, L, M, S, W or Q; XV_3 - K, R or T;

XV_4 - L, A or V; XV_5 - R, A, M, H, E, V, L, G, D, Q or S; XV_6 - E or

XV_7 - A or G; and XV_8 - R, H, G or I) and XVI_1XVI_2XVI_3XVI_4XVI_5

XVI_6XVI_7XVI_8XVI_9 (where XVI_1 - A, E or G; XVI_2 - E, H or D;

XVI_3 - R or G; XVI_4 - K, Y, M, N, Q, R, D, L, S or E; XVI_5 - A, S or

XVI_6 - E, D, T, O, K or A; XVI_7 - R, W, K, L, S, A or Q; XVI_8 - R

or E; and XVI_9 - W, G or R). The compounds are used for treating

conditions associated with depressed neutrophil count e.g. chemotherapy-

induced neutropenia, AIDS-induced neutropenia or community-acquired

pneumonia-induced pneumonia. The compounds are useful as in vitro as

tools for understanding the biological role of granulocyte-colony

stimulating factor (G-CSF) a haematopoietic growth factor and

cytokine that stimulates neutrophil proliferation and differentiation).

Including evaluation of many factors thought to influence, and be

influenced by, production of white blood cells, in the development of

compounds that bind to G-CSFR, as reagents for detecting G-CSF receptor

or related receptor on living cells, fixed cells, in biological fluid, in

tissue homogenates or in purified natural biological materials, in situ

staining, fluorescence-activated cell sorting (FACS), Western blotting or

enzyme-linked immunosorbent assay (ELISA), in receptor purification or

in purifying cells expressing G-CSFR on the cell surface (or inside

permeabilised cells) as a commercial research reagent for various medical

and diagnostic uses or to treat a disease that would benefit from the

ability to of a compound to mimic the effects of G-CSF in vivo.

The compounds bind specifically to G-CSFR and allow for studies of

biological activities mediated by the receptor and for the treatment of

diseases, disorders and conditions that would benefit from activating or

inactivating G-CSFR. The present sequence is a G-CSFR binding peptide of

the invention.

Query Match 2.4%; Score 6; DB 23; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ELSDSG 220

DB IIIII

6 ELSDSG 11

RESULT 30

AA89248

AA89248 standard; Peptide: 21 AA.

AA89248;

23-MAY-2001 (first entry)

```

XX HIV gp120 protein binding peptide #341.
XX
XX Human chemokine receptor; CD4; HIV; glycoprotein 120; gp120; antagonist;
KW replication; CCR5; CXCR4; CD4; STRL33.
XX
XX Homo sapiens.
XX
XX WC200116182-A2.
XX
XX 08-MAR-2001.
XX
XX 25 AUG-2000; 2000WO-US23505.
XX
XX 27-AUG-1999; 99US-0151270.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Saxinger C;
XX
XX WPI: 2001-244398/25.
XX
XX Novel polypeptides useful for treating HIV infection, have homology to
XX regions of domains of human chemokine receptors CCR5, CXCR4 and STRL33,
XX and binds to HIV gp120 under physiological conditions.
XX
XX Example 4; Page 46; 114pp; English.
XX
XX The present invention describes a number of peptides which are able to
XX bind to HIV glycoprotein 120 (gp120). These are similar to the human
XX chemokine receptors CCR5, CXCR4 and STRL33, as well as CD4. These are
XX useful in the treatment of HIV, as they prevent replication of the
XX virus. The present sequence is an example of a peptide of the invention.
XX
XX Sequence 21 AA:
XX
XX Query Match 2.4%; Score 6; DB 22; Length 21;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 216 LSDSGQ 221
XX
XX DB 12 LSDSGQ 17
XX
XX RESULT 31
XX AAB89245
XX
XX AC AAB89249;
XX
XX DT 23-MAY-2001 (first entry)
XX
XX XX HIV gp120 protein binding peptide #342.
XX
XX DE Human chemokine receptor; CD4; HIV; glycoprotein 120; gp120; antagonist;
XX KW replication; CCR5; CXCR4; CD4; STRL33.
XX
XX XX Homo sapiens.
XX
XX OS WC200116182-A2.
XX
XX FN 08-MAR-2001.
XX
XX PD 25-AUG-2000; 2000WO-US23505.
XX
XX PF 27-AUG-1999; 99US-0151270.
XX
XX PR (USSH ) US DEPT HEALTH & HUMAN SERV.CES.
XX
XX PA Saxinger C;
XX
XX P1 WPI: 2001-244398/25.
XX
XX DR

```

```

XX
XX Novel polypeptides useful for treating HIV infection, have homology to
XX regions of domains of human chemokine receptors CCR5, CXCR4 and STRL33,
XX and binds to HIV gp120 under physiological conditions.
XX
XX Example 4; Page 46; 114pp; English.
XX
XX The present invention describes a number of peptides which are able to
XX bind to HIV glycoprotein 120 (gp120). These are similar to the human
XX chemokine receptors CCR5, CXCR4 and STRL33, as well as CD4. These are
XX useful in the treatment of HIV, as they prevent replication of the
XX virus. The present sequence is an example of a peptide of the invention.
XX
XX Sequence 21 AA:
XX
XX Query Match 2.4%; Score 6; DB 22; Length 21;
XX Best Local Similarity 100.0%; Pred. No. 1.9e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 216 LSDSGQ 221
XX
XX DB 7 LSDSGQ 12
XX
XX RESULT 32
XX AAB89250
XX
XX TD AAB89250 standard; Peptide; 21 AA.
XX
XX AC AAB89250;
XX
XX XX 23-MAY-2001 (first entry)
XX
XX XX HIV gp120 protein binding peptide #343.
XX
XX KW Human chemokine receptor; CD4; HIV; glycoprotein 120; gp120; antagonist;
XX KW replication; CCR5; CXCR4; CD4; STRL33.
XX
XX OS Homo sapiens.
XX
XX PN WC200115182-A2.
XX
XX PD 08-MAR-2001.
XX
XX PF 25-AUG-2000; 2000WO-US23505.
XX
XX PR 27-AUG-1999; 99US-0151270.
XX
XX XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Saxinger C;
XX
XX DR WPI: 2001-244398/25.
XX
XX XX Novel polypeptides useful for treating HIV infection, have homology to
XX regions of domains of human chemokine receptors CCR5, CXCR4 and STRL33,
XX and binds to HIV gp120 under physiological conditions.
XX
XX Example 4; Page 46; 114pp; English.
XX
XX PS The present invention describes a number of peptides which are able to
XX bind to HIV glycoprotein 120 (gp120). These are similar to the human
XX chemokine receptors CCR5, CXCR4 and STRL33, as well as CD4. These are
XX useful in the treatment of HIV, as they prevent replication of the
XX virus. The present sequence is an example of a peptide of the invention.
XX
XX Sequence 21 AA:
XX
XX Query Match 2.4%; Score 6; DB 22; Length 21;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 216 LSDSGQ 221
XX
XX DB 11

```

```

Db      2 LSUSQ 7

RESULT 33
AAG78335
ID      AAG78335 standard; Peptide; 25 AA.
XX
XX
AC      AAG78935;
XX
XX
DT      13-FEB-2002 (first entry)
DE
DE
XX      Selfolobus solfataricus peptide fragment containing the Y-GGA/A motif.
XX
XX      Tag polymerase; B-type DNA polymerase.
XX
XX      Sulfolobus solfataricus.
XX
XX      EP1132474-A1.
XX
XX      12-SEP-2001.
XX
XX
XX      06-MAR-2001; 2001EP-0104583.
XX
XX      11-MAR-2000; 2000EP-0105155.
XX
XX      (HOFF ) ROCHER D:AGNOST:CS SMRH.
XX
XX      Sobek E, Frey B, Antraxikian G, Boelike K, Pisani FM, Rossi M;
XX      WPI: 2001-618367/72.
XX
XX      New thermostable mutant B-type DNA polymerase with a Y-GG/A amino acid
XX      motif between the N-terminal 3'-5' exonuclease domain and the
XX      C-terminal polymerase domain in the wild type polymerase, useful for
XX      synthesizing nucleic acids.
XX
XX      Disclosure; Fig 8; 40pp; English.
XX
XX
XX      The present invention relates to thermostable mutant B-type DNA
XX      polymerases, which have a Y-GG/A amino acid motif between the N-terminal
XX      3'-5' exonuclease domain and the C-terminal polymerase domain, where the
XX      tyrosine of this motif is mutated. The mutant B-type DNA polymerase is
XX      useful for synthesizing nucleic acids and for PCR. To illustrate the
XX      invention, thermococcus aggregans (Tag) DNA polymerase was used
XX      (AA170053). The present sequence is a peptide fragment of a B-type DNA
XX      polymerase, containing the Y-GGA/A motif, which was used in a sequence
XX      homology alignment.
XX
XX      Sequence 25 AA;
SQ
Query Match 2.4%; Score 6; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25 SAU:KG 30
DB      11-111

RESULT 34
ABG62532
ID      ABG62532 standard; Peptide; 25 AA.
XX
XX
AC      ABG62532;
XX
XX
DT      21-AUG-2002 (first entry)
DE
DE
XX      Eubacterial MutS1 DNA polymerase III: beta subunit binding peptide #37.
XX
XX      DNA polymerase III: beta subunit; eubacteria; antibacterial;
XX      eubacterial infection.
XX
XX      Ralstonia metallidurans.
XX

```

```

XX
XX      WO200238596-A1.
XX
XX      16-MAY-2002.
XX
XX      08-NOV-2001; 2001WO-AU01436.
XX
XX      08-NOV-2000; 2000AU-0001320.
XX      06-FEB-2001; 2001AU-0002919.
XX
XX      (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX
XX      Dairymple HP, Kongsuwan K, Wijffels GL, Jennings PA, Kemp GW;
XX      WPI: 2002-471546/50.
XX
XX      New molecule having surface analogous to surface of domain of
XX      eubacterial beta protein contacted by proteins that interact with beta
XX      protein, useful to identify inhibitors of beta protein-ligand
XX      interaction.
XX
XX      Example 1; Page 31; 326pp; English.
XX
XX      The invention relates to a molecule (I) comprising a surface (S)
XX      analogous to the surface of the domain of eubacterial beta protein
XX      contacted by proteins that interact with beta protein, where the
XX      surface is defined by the residues X(170), X(172), X(175), X(177),
XX      X(241), X(242), X(247), X(346), X(360), and X(362), where the
XX      superscript numbers designate the position of residues in Escherichia
XX      coli beta protein, or the equivalent residues in homologues from other
XX      species of eubacteria, and where:
XX      X(170) = Val, Ile, Ala, Thr, Ser or Glu; X(172) = Thr, Ser or Ile;
XX      X(175) = His, Tyr, Phe, Lys, Ile, Gln or Arg; X(177) = Leu, Met, Ile,
XX      Phe, Val or Ala; X(241) = Phe, Tyr or Leu; X(242) = Pro, Leu or Ile;
XX      X(247) = Val, Ile, Ala, Phe, Leu or Met; X(346) = Ser, Pro, Ala, Tyr or
XX      Lys; X(360) = Ile, Leu or Val; and X(362) = Met, Leu, Val, Ser, Thr or
XX      Arg. Also included are methods of identifying a modulator of the
XX      interaction between a eubacterial beta protein and proteins that interact
XX      with them, reducing (M4) the effect of eubacterial infestation of a
XX      biological system, involves delivering to a system infested with a
XX      eubacterial species, a modulator of the interaction between eubacterial
XX      beta protein and proteins that interact with the beta protein; and
XX      (4) a template (II) for the design of a compound that binds to at least
XX      part of (S) of beta protein as defined above comprises a (P) such as
XX      X'1X'2, X'3X'1X'2, X'3X'1X'2X'4, GlnX'5X'3X'1X'2, GlnX'5X'X'3X'6,
XX      where: x = any amino acid residue; X'1 = Leu, Met, Ile, or Phe;
XX      X'2 = Leu, Ile, Val, Cys, Phe, Tyr, Trp, Pro, Asp, Ala or Gly;
XX      X'3 = Ala, Gly, Thr, Asn, Asp, Ser, or Pro; X'4 = Ala or Gly; X'5 = Leu;
XX      and X'6 = Leu, Ile, Val, Cys, Phe, Tyr, Trp or Pro. The method are useful
XX      for identifying a modulator of the interaction between a eubacterial beta
XX      protein and proteins that interact with the beta protein. (M4) is useful
XX      for reducing the effect of eubacteria; infestation of a biological
XX      system. The compounds identified using above mentioned methods are
XX      useful as antibacterial agent for treatment or prevention of disease in
XX      humans, animals and plants. The present sequence is a eubacterial;
XX      peptide from a DNA binding protein or polymerase which contains a DNA
XX      polymerase III beta subunit binding site.
XX
XX      Sequence 25 AA;
SQ
Query Match 2.4%; Score 6; DB 23; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 MDLFS 26
DB      11111
        12 MDLFS 17

RESULT 35
AAG62760
ID      AAG62760 standard; peptide; 27 AA.
XX

```

AC AAG62760;
 DT 17-SEP-2001 (first entry)
 XX Amino acid sequence of target sequence from a hsp.
 DE
 XX Multi-helical lid; heat shock protein; hsp; protein folding;
 KW pathogenic infection; bacterial infection; antibacterial;
 XX Streptococcus pyogenes.
 OS
 XX WO200153509 A2.
 PN
 XX 26-JUL-2001.
 PD
 XX 19-JAN-2001; 2001WO-US01812.
 PF
 XX 21-JAN-2000; 2000US-0177565.
 PR 04-OCT-2000; 2000US-0237599.
 XX
 XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
 PA (UYCR-) UNIV CREIGHTON.
 XX
 XX Stivas L, Blaszczyk-Thurin M, Rogers M, Iovas S;
 PI WPI; 2001-451911/48.
 DR
 XX Composition, used to treat a pathogenic infection and eliminate a
 PT plant, insect, or animal pest, comprises a molecule that binds to a
 PP heat shock protein.
 XX
 XX Example 15; Page 86; 124pp; English.
 US
 XX The specification describes a composition that comprises a synthetic
 CC non naturally occurring molecule that binds to a selected multi-helical
 CC lid of a heat shock protein (hsp) of a selected organism, where the
 CC molecule inhibits protein folding activity of the hsp, and a carrier,
 CC where exposure of the organism to the composition retards the growth
 CC and reproduction of the organism. The composition is used to treat a
 CC mammal suffering from a pathogenic infection, in the manufacture of a
 CC medicament for treating a mammal for a pathogenic infection, and to
 CC eliminate a plant, insect, or animal pest. It is used in the manufacture
 CC of a medicament for treating mammalian bacterial infection. The present
 CC sequence represents a target sequence of a multi-helical lid of a hsp.
 XX
 XX Sequence 27 AA:
 SQ
 Query Match 2.4%; Score 6; DB 22; length 27;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 192 EKAAAL 197
 DB 10 EKAAAL 15
 |||||
 RESULT 36
 ANM18559
 ID ANM18559 standard; Protein: 29 AA.
 XX
 XX ANM18559;
 AC
 XX 12-OCT-2001 (first entry)
 DT
 XX Peptide #4993 encoded by probe for measuring cervical gene expression.
 DE
 KW Probe: human; microarray; gene expression; cervix; epithelial cell;
 KW cervical cancer.
 XX
 XX Homo sapiens.
 OS
 XX WO200157278 A2.
 PN
 XX

PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488901/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 PS
 XX Claim 27; SEQ ID No 23385; 487pp; English.
 XX
 XX The present invention relates to human single exon nucleic acid probes
 CC (SNP: see AAI0068-AA128459). The present sequence is a peptide encoded
 CC by one such probe. The SNPs are derived from human HeLa cells. The SNPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 29 AA:
 SQ
 Query Match 2.4%; Score 6; DB 22; length 29;
 Best Local Similarity 100.0%; Pred. No. 2.5e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 178 VNRKHS 183
 DB 14 VNRKHS 19
 |||||
 RESULT 37
 AAG62754
 ID AAG62754 standard; peptide; 33 AA.
 XX
 XX AAG62754;
 AC
 XX 17-SEP-2001 (first entry)
 DT
 XX Amino acid sequence of a D-E helix domain.
 DE
 XX Multi-helical lid; heat shock protein; hsp; protein folding;
 KW pathogenic infection; bacterial infection; antibacterial; Dnak.
 XX
 XX Streptococcus pyogenes.
 OS
 XX WO200153509-A2.
 PN
 XX 26-JUL-2001.
 PD
 XX 19-JAN-2001; 2001WO-US01812.
 PF
 XX 21-JAN-2000; 2000US-0177565.
 PR 03-OCT-2000; 2000US-0237599.
 XX
 XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.
 PA (UYCR-) UNIV CREIGHTON.
 XX

PI Olivos L, Blaszczyk-Thurin M, Rogers M, Lovas S;
 DR WPI: 2001 45191/48.
 XX
 XX Composition, used to treat a pathogenic infection and eliminate a
 PT plant, insect, or animal pest, comprises a molecule that binds to a
 PT heat shock protein.
 XX
 PS Claim 19; Page 9; 124pp; English.
 XX
 CC The specification describes a composition that comprises a synthetic
 CC non-naturally occurring molecule that binds to a selected multi-helical
 CC lid of a heat shock protein (hsp) of a selected organism, where the
 CC molecule inhibits protein folding activity of the hsp, and a carrier,
 CC where exposure of the organism to the composition retards the growth
 CC and reproduction of the organism. The composition is used to treat a
 CC mammal suffering from a pathogenic infection, in the manufacture of a
 CC medicament for treating a mammal for a pathogenic infection, and to
 CC eliminate a plant, insect, or animal pest. It is used in the manufacture
 CC of a medicament for treating mammalian bacterial infection. The present
 CC sequence represents a target sequence of a multi-helical lid of a hsp.
 XX
 SQ Sequence 33 AA;
 Query Match 2.4%; Score 6; DB 22; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 192 EKAQAL 197
 DB 13 EKAQAL 15
 IIII
 RESULT 34
 ID A8319418 standard; Protein; 38 AA.
 XX
 AC A8319418;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #19309.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical; imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN W0200175867-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-0508631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PT Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-9587; AAS835J5.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID No 49677; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. A830010-A830377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SQ Sequence 38 AA;
 Query Match 2.4%; Score 6; DB 22; Length 38;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 227 GGPAPG 232
 DB 10 GGPAPG 15
 IIII
 RESULT 39
 ID AAM65279 standard; Protein; 39 AA.
 XX
 AC AAM65279;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 37384.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN W0200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX
 PS Example 4; SEQ ID NO: 37384; 650pp - Sequence listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX Sequence 39 AA;

Query Match 2.4%; Score 6; DB 22; length 39;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 PKASSS 207

|||||

DB 7 PKASSS 12

RESULT 40

ABG46994

ID ARG46994 standard; Peptide: 39 AA.

XX AC ARG46994;

XX DT 19-AUG-2002 (first entry)

XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 36659.

XX KW Human; single exon probe; asthma; lung cancer; COPD; ILD;

XX KW chronic obstructive pulmonary disease; interstitial lung disease;

XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;

XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX KW primary ciliary dyskinesia; pulmonary hypertension;

XX KW hyaline membrane disease.

XX OS Homo sapiens.

XX PN W0200186003-A2.

XX PD 25-NOV-2001.

XX PF 30-JAN-2001; 2001W0-US00665.

XX PR 04-FEB-2000; 2000US-180312P.

XX PR 26-MAY-2000; 2000US-207456P.

XX PR 30-JUN-2000; 2000US-060840P.

XX PR 03-AUG-2000; 2000US-063236P.

XX PR 21-SEP-2000; 2000US-234687P.

XX PR 27-SEP-2000; 2000US-236359P.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MULE-) MOLECULAR DYNAMICS INC.

XX PT Penn SG, Hanzel DK, Chen W, Rank DK;

XX DR WPI; 2002-114183/15.

XX PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -

XX PS Claim 27; SEQ ID No 36659; 634pp; English.

XX CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a

CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 39 AA;

Query Match 2.4%; Score 6; DB 23; Length 39;

Best Local Similarity 100.0%; Pred. No. 3.2e-02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 PKASSS 207

|||||

DB 7 PKASSS 12

Search completed: April 10, 2003, 10:39:15

Job time : 31.1765 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: April 10, 2003, 10:29:45 ; Search time 13.5294 Seconds
(without alignments)
543.684 Million cell updates/sec

Title: US-09-930-026-1

Perfect score: 1315

Sequence: 1 MALLATS:PEGIMVKTFTEDR.....POEASQSDSBSGAGGLAPS 250

Scoring table: HLOSUMC2

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422322 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 135 summaries

Database : 1: /cun2.6/ptodata/1/iaa/5A_COMB.pep:*
2: /cun2.6/ptodata/1/iaa/5B_COMB.pep:*
3: /cun2.6/ptodata/1/iaa/6A_COMB.pep:*
4: /cun2.6/ptodata/1/iaa/6B_COMB.pep:*
5: /cun2.6/ptodata/1/iaa/PCIOUS_COMB.pep:*
6: /cun2.6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1315	100.0	250	4	US-09-058-368-1
2	179	13.6	172	2	US-08-318-947A-11
3	179	13.6	172	2	US-08-795-303-11
4	179	13.6	172	3	US-09-358-580-14
5	165.5	12.6	165	3	US-08-965-689A-6
6	165.5	12.6	165	4	US-09-359-967-6
7	164	12.5	197	2	US-08-949-289-1
8	164	12.5	197	4	US-08-328-866-1
9	159	12.1	151	2	US-08-318-947A-10
10	159	12.1	151	2	US-08-795-303-10
11	159	12.1	152	3	US-09-358-580-10
12	157	11.9	152	2	US-08-247-904B-16
13	157	11.9	152	3	US-08-767-942A-17
14	155	11.8	152	1	US-08-318-947A-7
15	155	11.8	152	2	US-08-795-303-7
16	154	11.7	152	2	US-08-318-947A-6
17	154	11.7	152	2	US-08-795-303-6
18	153	11.6	152	1	US-08-318-947A-8
19	153	11.6	152	2	US-08-795-303-8
20	152	11.6	152	3	US-09-358-580-8
21	151	11.5	151	1	US-08-318-947A-9
22	151	11.5	151	2	US-08-795-303-9
23	148	11.3	298	3	US-08-767-942A-25
24	148	11.3	298	4	US-09-177-165A-23
25	145.5	11.1	282	4	US-09-058-368-2
26	145	11.0	152	3	US-09-358-580-12
27	144.5	11.0	152	2	US-08-533-298-2

28	137.5	10.5	165	3	US-08-965-689A-3	Sequence 3, Appli
29	137.5	10.5	165	4	US-09-359-967-3	Sequence 3, Appli
30	134	10.2	147	2	US-08-464-604A-7	Sequence 7, Appli
31	132	10.0	148	3	US-09-358-580-2	Sequence 2, Appli
32	132	10.0	148	3	US-09-358-580-4	Sequence 4, Appli
33	131	10.0	148	3	US-09-358-580-6	Sequence 6, Appli
34	130	9.9	147	1	US-08-305-520-2	Sequence 2, Appli
35	130	9.9	147	2	US-08-989-289-3	Sequence 3, Appli
36	130	9.9	147	2	US-08-486-663A-2	Sequence 2, Appli
37	130	9.9	147	2	US-08-247-904B-2	Sequence 2, Appli
38	130	9.9	147	3	US-08-895-601-8	Sequence 8, Appli
39	130	9.9	147	3	US-08-767-942A-2	Sequence 2, Appli
40	130	9.9	147	4	US-09-328-866-3	Sequence 3, Appli
41	128.5	9.8	178	4	US-08-828-533-3	Sequence 3, Appli
42	128	9.7	147	1	US-08-305-520-6	Sequence 6, Appli
43	128	9.7	147	2	US-08-486-663A-6	Sequence 6, Appli
44	128	9.7	147	3	US-08-767-942A-6	Sequence 6, Appli
45	127	9.7	148	2	US-08-679-765-4	Sequence 4, Appli
46	127	9.7	148	2	US-08-464-604A-9	Sequence 9, Appli
47	127	9.7	148	2	US-09-196-525-4	Sequence 4, Appli
48	127	9.7	148	4	US-09-318-317-4	Sequence 4, Appli
49	127	9.7	157	1	US-08-706-214-4	Sequence 4, Appli
50	127	9.7	157	3	US-08-350-468-6	Sequence 6, Appli
51	126.5	9.6	164	2	US-08-918-723-3	Sequence 3, Appli
52	126.5	9.6	164	2	US-09-237-507-3	Sequence 3, Appli
53	124	9.4	147	1	US-08-305-520-4	Sequence 4, Appli
54	124	9.4	147	2	US-08-486-663A-4	Sequence 4, Appli
55	124	9.4	147	3	US-08-767-942A-4	Sequence 4, Appli
56	121	9.2	147	2	US-08-748-703-4	Sequence 4, Appli
57	121	9.2	147	3	US-09-132-861-4	Sequence 4, Appli
58	119	9.0	158	1	US-08-399-696-2	Sequence 2, Appli
59	119	9.0	250	2	US-08-861-269-5	Sequence 5, Appli
60	119	9.0	250	2	US-09-134-596-5	Sequence 5, Appli
61	119	9.0	250	3	US-09-293-273-5	Sequence 5, Appli
62	118.5	9.0	166	2	US-08-918-723-4	Sequence 4, Appli
63	118.5	9.0	166	2	US-09-237-507-4	Sequence 4, Appli
64	118.5	9.0	177	4	US-08-828-533-8	Sequence 8, Appli
65	118	9.0	142	3	US-08-350-468-2	Sequence 2, Appli
66	118	9.0	147	1	US-08-305-520-7	Sequence 7, Appli
67	118	9.0	147	2	US-08-464-604A-2	Sequence 2, Appli
68	118	9.0	147	2	US-08-486-663A-7	Sequence 7, Appli
69	118	9.0	147	3	US-08-767-942A-7	Sequence 7, Appli
70	118	9.0	158	1	US-08-706-214-3	Sequence 3, Appli
71	118	9.0	158	1	US-08-318-947A-2	Sequence 2, Appli
72	118	9.0	158	2	US-08-795-303-2	Sequence 2, Appli
73	118	9.0	158	2	US-08-486-663A-13	Sequence 13, Appli
74	118	9.0	158	3	US-08-767-942A-13	Sequence 13, Appli
75	118	9.0	158	3	US-08-350-468-4	Sequence 4, Appli
76	118	9.0	158	5	PCT-US95-06722-24	Sequence 24, Appli
77	117	8.9	158	1	US-08-318-947A-5	Sequence 5, Appli
78	117	8.9	158	2	US-08-795-303-5	Sequence 5, Appli
79	117	8.9	170	2	US-08-820-170A-22	Sequence 22, Appli
80	117	8.9	170	2	US-08-918-723-1	Sequence 1, Appli
81	117	8.9	170	2	US-09-237-507-1	Sequence 1, Appli
82	117	8.9	170	3	US-09-055-699-22	Sequence 22, Appli
83	117	8.9	170	4	US-09-273-565-22	Sequence 22, Appli
84	117	8.9	170	4	US-09-565-538-22	Sequence 22, Appli
85	117	8.9	170	4	US-09-661-468-22	Sequence 22, Appli
86	116.5	8.9	179	4	US-08-828-533-1	Sequence 1, Appli
87	116	8.8	147	1	US-08-464-342-2	Sequence 2, Appli
88	116	8.8	147	2	US-08-875-272-2	Sequence 2, Appli
89	116	8.8	147	2	US-08-903-396-2	Sequence 2, Appli
90	110.5	8.4	253	1	US-08-399-696-4	Sequence 4, Appli
91	110.5	8.4	253	1	US-08-399-696-118	Sequence 118, App
92	109	8.3	295	2	US-08-679-765-5	Sequence 5, Appli
93	109	8.3	295	2	US-09-196-525-5	Sequence 5, Appli
94	109	8.3	295	4	US-09-318-317-5	Sequence 5, Appli
95	109	8.3	295	4	US-09-177-165A-22	Sequence 22, Appli
96	107.5	8.2	154	1	US-08-464-342-4	Sequence 4, Appli
97	107.5	8.2	154	2	US-08-748-703-3	Sequence 3, Appli
98	107.5	8.2	154	2	US-08-464-604A-4	Sequence 4, Appli
99	107.5	8.2	154	2	US-08-875-272-4	Sequence 4, Appli
100	107.5	8.2	154	2	US-08-718-538-5	Sequence 5, Appli

101 107.5 8.2 154 2 US-08-903-396-4 Sequence 4, Appli
102 107.5 8.2 154 3 US-09-132-661-3 Sequence 3, Appli
103 106.5 8.1 154 2 US-09-132-661-3 Sequence 3, Appli
104 103 7.8 180 4 US-08-828-533-6 Sequence 6, Appli
105 100.5 7.6 162 2 US-08-464-604A-8 Sequence 8, Appli
106 99.5 7.6 193 1 US-08-464-342-6 Sequence 6, Appli
107 99.5 7.6 193 2 US-08-875-272-6 Sequence 6, Appli
108 99.5 7.6 193 2 US-08-903-396-6 Sequence 6, Appli
109 98 7.5 540 1 US-07-906-930E-2 Sequence 2, Appli
110 96 7.3 153 2 US-08-748-703-1 Sequence 1, Appli
111 96 7.3 153 3 US-09-132-861-3 Sequence 1, Appli
112 96 7.3 494 1 US-07-906-930E-33 Sequence 33, Appli
113 95.5 7.3 318 2 US-08-960-022-4 Sequence 4, Appli
114 94.5 7.2 201 2 US-08-679-765-1 Sequence 1, Appli
115 94.5 7.2 201 2 US-03-396-525-1 Sequence 1, Appli
116 94.5 7.2 201 4 US-09-318-317-1 Sequence 1, Appli
117 94.5 7.2 259 2 US-08-861-269-1 Sequence 1, Appli
118 94.5 7.2 259 2 US-09-334-596-1 Sequence 1, Appli
119 94.5 7.2 259 3 US-09-293-273-1 Sequence 1, Appli
120 94 7.1 185 2 US-08-933-750C-2 Sequence 2, Appli
121 94 7.1 185 3 US-08-965-689A-1 Sequence 1, Appli
122 94 7.1 185 4 US-09-234-613-2 Sequence 2, Appli
123 94 7.1 185 4 US-09-359-967-1 Sequence 1, Appli
124 93.5 7.1 318 4 US-09-058-368-3 Sequence 3, Appli
125 92.5 7.0 119 3 US-09-012-518A-24 Sequence 24, Appli
126 92.5 7.0 119 4 US-08-360-144A-24 Sequence 24, Appli
127 92.5 7.0 193 2 US-08-679-765-3 Sequence 3, Appli
128 92.5 7.0 193 2 US-08-464-604A-6 Sequence 6, Appli
129 92.5 7.0 193 2 US-09-196-525-3 Sequence 3, Appli
130 92.5 7.0 193 4 US-09-318-317-3 Sequence 3, Appli
131 86 6.5 462 2 US-08-865-597A-2 Sequence 2, Appli
132 83 6.3 434 4 US-09-012-504A-24 Sequence 24, Appli
133 82.5 6.3 434 4 US-08-630-915A-22 Sequence 22, Appli
134 82 6.2 349 3 US-08-469-318-151 Sequence 151, App
135 82 6.2 349 3 US-08-468-609A-151 Sequence 151, App

ALIGNMENTS

RESULT 1
US-09-058-368-1
: Sequence 1, Application US/09058368
: Patent No. 6277568
: GENERAL INFORMATION:
: APPLICANT: Lal, Prateet
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Guebler, Karl T.
: APPLICANT: Corley, Neil C.
: APPLICANT: Baughn, Mariah R.
: APPLICANT: Azimzai, Yalda
: TITLE OF INVENTION: HUMAN UBIQUITIN-CONJUGATING ENZYME
: TITLE OF INVENTION: HOMOLOGS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/058,368
: FILING DATE: HEREWITH
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:

ATTORNEY/AGENT INFORMATION:
: NAME: Cerrione, Michael C
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PR-0502 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 250 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: PROSNOT14
: CLONE: 1728211
: US-09-058-368-1
Query Match 100.0%; Score 1315; DB 4; Length 250;
Best Local Similarity 100.0%; Pred. No. 2e-137;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALLATSLPGLIMVKTFFDRMDLFSALIKGPTPTPTDGLYLFDIQLPNYPVPPHFCY 60
|||||
Db 1 MALLATSLPGLIMVKTFFDRMDLFSALIKGPTPTPTDGLYLFDIQLPNYPVPPHFCY 60
QY 61 LSQCSGRNLNPLYDNGKVCVSLGTWIGKGTERTWTSKSLLOVLISIOGLILVNEPYNFR 120
|||||
Db 61 LSQCSGRNLNPLYDNGKVCVSLGTWIGKGTERTWTSKSLLOVLISIOGLILVNEPYNFR 120
QY 121 AGFSDSRLQEGYENSRVCYNEMALIRVVQSMTOLVRRPPEVFEQEIQRHSTGGWRLVNR 180
|||||
Db 121 AGFSDSRLQEGYENSRVCYNEMALIRVVQSMTOLVRRPPEVFEQEIQRHSTGGWRLVNR 180
QY 181 IESWLETHALLEKAQALPNVGPVKASSPPPPAVPAELSDSGQEPEDGPGAPGASQSGDS 240
|||||
Db 181 IESWLETHALLEKAQALPNVGPVKASSPPPPAVPAELSDSGQEPEDGPGAPGASQSGDS 240
QY 241 EGGAQGLAFS 250
|||||
Db 241 EGGAQGLAFS 250

RESULT 2

US-08-318-947A-11
: Sequence 11, Application US/08318947A
: Patent No. 578245
: GENERAL INFORMATION:
: APPLICANT: Anderson, Paul J.
: APPLICANT: Tian, Qingsheng
: TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
: TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sughrie, Mion, Zinn, Macpeak & Seas
: STREET: 2100 Pennsylvania Avenue, NW Suite 800
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/318,947A
: FILING DATE: 06-OCT-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/133,530
: FILING DATE: 07-OCT-1993


```

; INDEX: 0491103
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 172 amino acids
; TYPE: amino acid

```

```

;
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Diskette
;
; COMPUTER: IBM Compatible
;
; OPERATING SYSTEM: DOS
;
; SOFTWARE: FastSEQ for Windows Version 2.0
;

```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/965,689A
: FILING DATE: Herewith
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/933,750
: FILING DATE: September 23, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356-1 CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 165 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 4257
: US-08-965-689A-6

Query Match 12.6%; Score 165.5; DB 3; Length 165;
Best Local Similarity 31.2%; Pred. No. 2e-10;
Matches 39; Conservative 28; Mismatches 45; Indels 13; Gaps 67

QY 9 PPGIMV-KTFEDMDLFSALIKGPTPTPYEGLYLFDIQ:PNLIYPAVPPHFCYI:SQSCGR 67
DB 21 PPGIVAGPKSENNIFWDGLIQGPDPTFYADGVFNAKLEFPKDYPLSPKLTFTPTST--- 77
QY 68 LNPNIYONGKVCVSLIGT-----WIGKGTERTSKSSLQVLISQGLIIVNEPYNEA 121
DB 78 LHPNIYPNGVCIS:LHSQDDPNMYELAEERWSPVQSVKILLVSMML--SEPNI-ES 134
QY 122 GFDSQ 126
DB 135 GANID 139

RESULT 6
US-09-359-967-6
: Sequence 6, Application US/09459967
: Patent No. 6346624
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN UBICUITIN CONJUGATING ENZYMES
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/355,967
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/965,689
: FILING DATE:
: APPLICATION NUMBER: 08/933,750
```

```

: FILING DATE: September 23, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356-1 CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 165 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 4257
: US-09-359-967-6

Query Match 12.6%; Score 165.5; DB 4; Length 165;
Best Local Similarity 31.2%; Pred. No. 2e-10;
Matches 39; Conservative 28; Mismatches 45; Indels 13; Gaps 5;

QY 9 PPGIMV-KTFEDMDLFSALIKGPTPTPYEGLYLFDIQ:PNLIYPAVPPHFCYI:SQSCGR 67
DB 21 PPGIVAGPKSENNIFWDGLIQGPDPTFYADGVFNAKLEFPKDYPLSPKLTFTPTST--- 77
QY 68 LNPNIYONGKVCVSLIGT-----WIGKGTERTSKSSLQVLISQGLIIVNEPYNEA 121
DB 78 LHPNIYPNGVCIS:LHSQDDPNMYELAEERWSPVQSVKILLVSMML--SEPNI-ES 134
QY 122 GFDSQ 126
DB 135 GANID 139

RESULT 7
US-08-989-289-1
: Sequence 1, Application US/08989289
: Patent No. 5968747
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Shah, Purvi
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: UBICUITIN-LIKE CONJUGATING PROTEIN
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/989,289
: FILING DATE: Herewith
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0439 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
```

TELEX:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 197 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: ADREU105
 CLONE: 2501808
 US-08-930-289-1

Query Match 12.5%: Score 164; DB 2; Length 197;
 Best Local Similarity 28.3%; Pred. No. 3.8e-10;
 Matches 49; Conservative 33; Mismatches 75; Indels 16; Gaps 6;

QY 3 LIA-SLEGIWVTFEDRMOLFSALIKGPTPTPYDGLYLFDTQLPNTIYPAVPPHFCYLS 62
 DB 13 MLATEPPGLICWQDKQDDMLRAQLGGANTPYEKGVEKLEVIIPERYPEPPIRELT 72
 QY 64 QCSGRLEPNLYDNGKVCVSLGTWIGKTERWTSKSLQVLISIQGLILVNEPYYNEAG 122
 DB 73 PI---YHPNIDSAGRICLDVL-KLPPKGA--WRPSLNTATVLSIQ--LLMSEP----- 118
 QY 123 FDSORGLQECYENSCYNEMLIRVQSMTOL-VRRPPEVEFQETROHFSTGG 174
 DB 119 -NPDDPLMADISSEFYKYPAPLKNAROWTEKHAKQKQADEEMLDNLPEAG 170

RESULT 6
 US-09-328-866-1
 Sequence 1, Application US/09328866
 Patent No. 635358
 GENERAL INFORMATION:
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Shah, Puvi G.
 APPLICANT: Corley, Neil G.
 TITLE OF INVENTION: UBIQUITIN-LIKE CONJUGATING PROTEIN
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09328.866
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/989,289
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PP-0439 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650 855-0555
 TELEFAX: 650-845-4166
 TELEX:

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 197 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:

LIBRARY: ADREU105
 CLONE: 2501808
 US-09-328-866-1

Query Match 12.5%: Score 164; DB 4; Length 197;
 Best Local Similarity 28.3%; Pred. No. 3.8e-10;
 Matches 49; Conservative 33; Mismatches 75; Indels 16; Gaps 6;

QY 3 LIA-SLEGIWVTFEDRMOLFSALIKGPTPTPYDGLYLFDTQLPNTIYPAVPPHFCYLS 62
 DB 13 MLATEPPGLICWQDKQDDMLRAQLGGANTPYEKGVEKLEVIIPERYPEPPIRELT 72
 QY 63 QCSGRLEPNLYDNGKVCVSLGTWIGKTERWTSKSLQVLISIQGLILVNEPYYNEAG 122
 DB 73 PI---YHPNIDSAGRICLDVL-KLPPKGA--WRPSLNTATVLSIQ--LLMSEP----- 118
 QY 123 FDSORGLQECYENSCYNEMLIRVQSMTOL-VRRPPEVEFQETROHFSTGG 174
 DB 119 -NPDDPLMADISSEFYKYPAPLKNAROWTEKHAKQKQADEEMLDNLPEAG 170

RESULT 9
 US-08-318-947A-10
 Sequence 10, Application US/08318947A
 Patent No. 5798245
 GENERAL INFORMATION:
 APPLICANT: Anderson, Paul J.
 APPLICANT: Tian, Qingsheng
 TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
 TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sughrie, Mion, Zinn, Macpeak & Seas
 STREET: 2100 Pennsylvania Avenue, NW Suite 800
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20037

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/318,947A
 FILING DATE: 06-OCT-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/133,530
 FILING DATE: 07-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Mack, Susan J.
 REGISTRATION NUMBER: 30,951
 REFERENCE/DOCKET NUMBER: A6462
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)293-7060
 TELEFAX: (202)293-2920
 TELEX: 6491103

INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 151 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-318-947A-10

Query Match 12.1%: Score 159; DB 1; Length 151;
 Best Local Similarity 30.5%; Pred. No. 9e-10;
 Matches 40; Conservative 23; Mismatches 54; Indels 14; Gaps 4;

QY 9 PEGIMVTFEDRMOLFSALIKGPTPTPYDGLYLFDTQLPNTIYPAVPPHFCYLSQCSGRL 68
 DB 119 -NPDDPLMADISSEFYKYPAPLKNAROWTEKHAKQKQADEEMLDNLPEAG 170

DB 21 PAGVSASIVSDNVMWLNVAITDPACTFFEDGTFKWLSPDEQYPNKPPVWKFVST---MF 77
 QY 69 NPNLYDNGKVCVSLIGTWIGKTERWTSKSSLLQVLISIQGLIIVNPPYNEAGFSDRG 128
 DB 78 HPNVYANGELCLDIL-----QNRWSTYDVAAIITSQSL--NDP--NNASPANAEA 126
 QY 129 LQGYENSRCY 139
 DB 127 AQLHRENKKEY 137

RESULT 10
 US-08-795-303-10
 ; Sequence 10, Application US/08795303
 ; Patent No. 5948556
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, Paul J.
 ; APPLICANT: Tian, Qingsheng
 ; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
 ; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Suphache, Miao, Zinn, Macpeak & Seas
 ; STREET: 2108 Pennsylvania Avenue, NW Suite 800
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-PCS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/795,303
 ; FILING DATE: 04-FEB-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/318,947
 ; FILING DATE: 06-OCT-1994
 ; APPLICATION NUMBER: 08/133,530
 ; FILING DATE: 07-OCT-1993
 ; ATTORNEY/AGNT INFORMATION:
 ; NAME: Mack, Susan J.
 ; REGISTRATION NUMBER: 30,951
 ; REFERENCE/DOCKET NUMBER: A6462
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)293-7660
 ; TELEFAX: (202)293-2920
 ; TELEX: 6491103
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 151 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-795-303-10

Query Match 32.18; Score 159; DB 2; Length 151;
 Best Local Similarity 30.58; Prod. No. 9c-10;
 Matches 40; Conservative 23; Mismatches 54; Indels 14; Gaps 4;

QY 9 PEGIMVKTFEDRMDFLSALIKGPTRTPYEDGLYLFDIQLPNIPYAVPHFCYLSQCSGRL 68
 DB 21 PAGVSASIVSDNVMWLNVAITDPACTFFEDGTFKWLSPDEQYPNKPPVWKFVST---MF 77
 QY 69 NPNLYDNGKVCVSLIGTWIGKTERWTSKSSLLQVLISIQGLIIVNPPYNEAGFSDRG 128
 DB 78 HPNVYANGELCLDIL-----QNRWSTYDVAAIITSQSL--NDP--NNASPANAEA 126
 QY 129 LQGYENSRCY 139
 DB 127 AQLHRENKKEY 137

DB 127 AQLHRENKKEY 137

RESULT 11
 US-09-358-580-10
 ; Sequence 10, Application US/09358580
 ; Patent No. 6107545
 ; GENERAL INFORMATION:
 ; APPLICANT: Mahajan, Pramod B.
 ; TITLE OF INVENTION: Maize Rad6 Genes and Uses Thereof
 ; FILE REFERENCE: 0883
 ; CURRENT APPLICATION NUMBER: US/09/358,580
 ; CURRENT FILING DATE: 1999-07-21
 ; EARLIER APPLICATION NUMBER: 60/096,546
 ; EARLIER FILING DATE: 1998-08-14
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 152
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; US-09-358-580-10

Query Match 12.18; Score 159; DB 3; Length 152;
 Best Local Similarity 29.38; Prod. No. 9e-10;
 Matches 43; Conservative 27; Mismatches 57; Indels 20; Gaps 4;

QY 9 PEGIMVKTFEDRMDFLSALIKGPTRTPYEDGLYLFDIQLPNIPYAVPHFCYLSQCSGRL 68
 DB 21 PAGVSASIVSDNVMWLNVAITDPACTFFEDGTFKWLSPDEQYPNKPPVWKFVST---MF 77
 QY 69 NPNLYDNGKVCVSLIGTWIGKTERWTSKSSLLQVLISIQGLIIVNPPYNEAGFSDRG 125
 DB 78 HPNVYANGELCLDIL-----QNRWSTYDVAAIITSQSL--NDP--NNASPANAEA 127
 QY 126 DRGLQECYENSRCYNEMALIKVQSMF 152
 DB 128 -----RMFSENKREYNKRYKRVVEQSWT 150

RESULT 12
 US-08-247-904B-16
 ; Sequence 16, Application US/08247904B
 ; Patent No. 5981699
 ; GENERAL INFORMATION:
 ; APPLICANT: Rolfe, Mark
 ; APPLICANT: Eckstein, Jens W.
 ; APPLICANT: Draetta, Giulio
 ; TITLE OF INVENTION: Human Ubiquitin: Conjugating Enzyme
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley, Hoag & Elliot
 ; STREET: One Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/247,904B
 ; FILING DATE: 23-MAY-1994
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGNT INFORMATION:
 ; NAME: Vincent, Matthew P.
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MIV-029.01
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 832-1000
 ; TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 152 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08 247-904B-16

Query Match 11.94; Score 157; DB 2; Length 152;

Best Local Similarity 30.38; Pred. No. 1.5e-09;
 Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;

QY 9 PEGIMVKTFEDRMDLFSALIKGPTFTFYEDGLYFDIQLPNLYPAVPPHFCYLSQCSGRL 68
 DB 21 PEGVSGAPSENNIMQWNAVIFGPGTFFEDGTFKLVIFEESEYPNKPPTVRELKSK---MF 77
 QY 69 NPPLYDNGKVCVSLIGTWIGKTERWTSKSSLLQVLISIQGLILVNEPYYNEAGFDSDRG 128
 DB 78 HPNYADGSLICDLIL-----QNRWSPYDVSSILTSIQSL--DEPNPNSPA--NSQA 126
 QY 129 LQEGYENSCYCNEMALIRVWVS 150
 DB 127 AOLYQENKREYKRVSAIVEQS 148

RESULT 13
 US-08-767-942A-17
 Sequence 17, Application US/08/767942A
 Patent No. 6068982
 GENERAL INFORMATION:
 APPLICANT: Rolfe, Mark
 APPLICANT: Chiu, M. Tsabei
 APPLICANT: Berlin, Vivian
 APPLICANT: Damagrez, Veronique
 APPLICANT: Draetta, Giulio
 APPLICANT: Guillaume, Sotarel
 TITLE OF INVENTION: GHIQUITIN CONJUGATING ENZYMES
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESS: FOLEY, ROAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/767,942A
 FILING DATE: 17-DEC-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MIV-029, 04
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 152 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-767-942A-17

Query Match 11.94; Score 157; DB 3; Length 152;

Best Local Similarity 30.38; Pred. No. 1.5e-09;
 Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;

QY 9 PEGIMVKTFEDRMDLFSALIKGPTFTFYEDGLYFDIQLPNLYPAVPPHFCYLSQCSGRL 68

DB 21 PEGVSGAPSENNIMQWNAVIFGPGTFFEDGTFKLVIFEESEYPNKPPTVRELKSK---MF 77
 QY 69 NPPLYDNGKVCVSLIGTWIGKTERWTSKSSLLQVLISIQGLILVNEPYYNEAGFDSDRG 128
 DB 78 HPNYADGSLICDLIL-----QNRWSPYDVSSILTSIQSL--DEPNPNSPA--NSQA 126
 QY 129 LQEGYENSCYCNEMALIRVWVS 150
 DB 127 AOLYQENKREYKRVSAIVEQS 148

RESULT 14
 US-08-318-947A-7
 Sequence 7, Application US/08318947A
 Patent No. 5798245
 GENERAL INFORMATION:
 APPLICANT: Anderson, Paul J.
 APPLICANT: Tian, Qingsheng
 TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
 TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
 STREET: 2100 Pennsylvania Avenue, NW Suite 800
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/318,947A
 FILING DATE: 06-OCT-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/133,530
 FILING DATE: 07-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Mack, Susan J.
 REGISTRATION NUMBER: 30,951
 REFERENCE/DOCKET NUMBER: A6462
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)293-7060
 TELEFAX: (202)293-2920
 TELEX: 6491103
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 152 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-318-947A-7

Query Match 11.88; Score 155; DB 1; Length 152;

Best Local Similarity 30.38; Pred. No. 2.5e-09;

Matches 43; Conservative 24; Mismatches 61; Indels 14; Gaps 4;

QY 9 PEGIMVKTFEDRMDLFSALIKGPTFTFYEDGLYFDIQLPNLYPAVPPHFCYLSQCSGRL 68
 DB 21 PEGVSGAPSENNIMQWNAVIFGPGTFFEDGTFKLVIFEESEYPNKPPTVRELKSK---MF 77
 QY 69 NPPLYDNGKVCVSLIGTWIGKTERWTSKSSLLQVLISIQGLILVNEPYYNEAGFDSDRG 128
 DB 78 HPNYADGSLICDLIL-----QNRWSPYDVSSILTSIQSL--DEPNPNSPA--NSQA 126
 QY 129 LQEGYENSCYCNEMALIRVWVS 150
 DB 127 AOLYQENKREYKRVSAIVEQS 148


```

: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/795,303
: FILING DATE: 04-FEB-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/318,947
: FILING DATE: 06-OCT-1994
: APPLICATION NUMBER: 08/133,530
: FILING DATE: 07-OCT-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Mack, Susan J.
: REGISTRATION NUMBER: 30,951
: REFERENCE/DOCKET NUMBER: A6462
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060
: TELEFAX: (202)293-2920
: TELEX: 6491103
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 152 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-795-303-6

Query Match 11.78; Score 154; DB 2; Length 152;
Best Local Similarity 30.38; Pred. No. 3.2e-09;
Matches 43; Conservative 24; Mismatches 6; Indels 14; Gaps 4;

QY 9 PEGIMVKTFEDRMDFSAIKGPTPTPYEDGLYFDIQLPNTYPAVPPHFCYLSQCSGRL 68
DB 21 PAVSCAUSENNIMVNAVIFGPGTPEGDTFKLLIEETEEYPNKPTVRFVSK---MF 77
QY 69 NPNLYDNCKVCSLLGTWIGKTERWTSKSLQLVLISIQGLIIVNEPYNNAAGFDSRG 128
DB 78 HPNVAGSGICDIL-----QNRWSPTYDVSSILTSIQSLC--EPNPNSPA--NSQA 126
QY 129 IQGYENSRNCYNEMALIRVQVS 150
DB 127 AQIQENKREYKRVSAIVEQS 148

RESULT 18
US-08-318-947A-B
: Sequence 8, Application US/083:8947A
: Patent No. 5798245
: GENERAL INFORMATION:
: APPLICANT: Anderson, Paul J.
: TITLE OF INVENTION: IIA-1 BINDING PROTEINS AND ISOLATED
: TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
: STREET: 2100 Pennsylvania Avenue, NW Suite 800
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/318,947A

```

```

: FILING DATE: 06-OCT-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/133,530
: FILING DATE: 07-OCT-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Mack, Susan J.
: REGISTRATION NUMBER: 30,951
: REFERENCE/DOCKET NUMBER: A6462
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060
: TELEFAX: (202)293-2920
: TELEX: 6491103
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 152 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-318-947A-8

Query Match 11.68; Score 153; DB 1; Length 152;
Best Local Similarity 28.98; Pred. No. 4.2e-09;
Matches 41; Conservative 27; Mismatches 60; Indels 14; Gaps 4;

QY 9 PEGIMVKTFEDRMDFSAIKGPTPTPYEDGLYFDIQLPNTYPAVPPHFCYLSQCSGRL 68
DB 21 PAVSCAUSENNIMVNAVIFGPGTPEGDTFKLLIEETEEYPNKPTVRFVSK---MF 77
QY 69 NPNLYDNCKVCSLLGTWIGKTERWTSKSLQLVLISIQGLIIVNEPYNNAAGFDSRG 128
DB 78 HPNVAGSGICDIL-----QNRWSPTYDVSSILTSIQSLC--EPNPNSPA--NSQA 126
QY 129 IQGYENSRNCYNEMALIRVQVS 150
DB 127 AQIQENKREYKRVSAIVEQS 148

RESULT 19
US-08-795-303-8
: Sequence 8, Application US/08795303
: Patent No. 5948656
: GENERAL INFORMATION:
: APPLICANT: Anderson, Paul J.
: TITLE OF INVENTION: IIA-1 BINDING PROTEINS AND ISOLATED
: TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
: STREET: 2100 Pennsylvania Avenue, NW Suite 800
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/795,303
: FILING DATE: 04-FEB-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/318,947
: FILING DATE: 06-OCT-1994
: APPLICATION NUMBER: 08/133,530
: FILING DATE: 07-OCT-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Mack, Susan J.
: REGISTRATION NUMBER: 30,951

```

REFERENCE/DOCKET NUMBER: A6462
TELEPHONE: (202)293-7960
TELEFAX: (202)293-2920
TELEX: 6491163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-795-303-8

Query Match 11.6%; Score 153; DB 2; Length 152;
Best Local Similarity 28.9%; Pred. No. 4.2e-09;
Matches 41; Conservative 27; Mismatches 60; Indels 14; Gaps 4;

QY 9 PEGIMVKTFFDRMDFSLIKGPTPTPYEDGGLYFDIQLPNIVAVPAPHCYLSQCSGRL 68
DB 21 PAVGSGAPTSNNIMWNAVIFGPHDTPEDGFKLTIEETEPYKPKPTVRFVSKV---MF 77
QY 69 NPNLYDNCKVCVSLGGLTWIGKGTERTWTSKSLLOVLISIGLLILVNEPYNEAGFDSRG 128
DB 78 HPNIYVAGSGICLDL-----QNRWSPRYDVSAIILTSIQSLI--SDPNPNSPANSTAAQ 128
QY 129 LOEGYENSRCYNEMALIRVQVS 150
DB 127 AOLYCNKREYKRVSAI VQVS 148

RESULT 20
US-09-358-580-8
Sequence 8, Application US/09358580
Patent No. 6:07545
GENERAL INFORMATION:
APPLICANT: Mahajan, Prasad B.
TITLE OF INVENTION: Maize Rad6 Genes and Uses Thereof
FILE REFERENCE: 0883
CURRENT APPLICATION NUMBER: US/09/358,580
CURRENT FILING DATE: 1999-07-21
EARLIER APPLICATION NUMBER: 62/096,546
EARLIER FILING DATE: 1998-08-14
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 152
TYPE: PRT
ORGANISM: Zea mays
US-09-358-580-8

Query Match 11.6%; Score 152; DB 3; Length 152;
Best Local Similarity 28.6%; Pred. No. 5.4e-09;
Matches 42; Conservative 27; Mismatches 58; Indels 20; Gaps 4;

QY 9 PEGIMVKTFFDRMDFSLIKGPTPTPYEDGGLYFDIQLPNIVAVPAPHCYLSQCSGRL 68
DB 21 PAVGSGAPTSNNIMWNAVIFGPHDTPEDGFKLTIEETEPYKPKPTVRFVSKV---MF 77
QY 69 NPNLYDNCKVCVSLGGLTWIGKGTERTWTSKSLLOVLISIGLLILVNEPYNEAGFDS 125
DB 78 HPNIYVAGSGICLDL-----QNRWSPRYDVSAIILTSIQSLI--SDPNPNSPANSTAAQ 128
QY 126 DRLOQGYENSRCYNEMALIRVQVSM 152
DB 128 ---RMSEKREYKRVVEVQSM 150

RESULT 21
US-08-318-947A-9
Sequence 9, Application US/08318947A
Patent No. 5798245
GENERAL INFORMATION:
APPLICANT: Anderson, Paul J.
TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, NW Suite 800
CITY: Washington
STATE: DC

APPLICANT: Anderson, Paul J.
TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, NW Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,947A
FILING DATE: 06-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/133,530
FILING DATE: 07-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951
REFERENCE/DOCKET NUMBER: A6462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-2920
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-318-947A-9

Query Match 11.5%; Score 151; DB 1; Length 151;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 40; Conservative 30; Mismatches 58; Indels 14; Gaps 4;

QY 9 PEGIMVKTFFDRMDFSLIKGPTPTPYEDGGLYFDIQLPNIVAVPAPHCYLSQCSGRL 68
DB 21 PAVGSGAPTSNNIMWNAVIFGPHDTPEDGFKLTIEETEPYKPKPTVRFVSKV---F 77
QY 69 NPNLYDNCKVCVSLGGLTWIGKGTERTWTSKSLLOVLISIGLLILVNEPYNEAGFDSRG 128
DB 78 HPNIYVAGSGICLDL-----QNRWSPRYDVSAIILTSIQSLI--SDPNPNSPANSTAAQ 128
QY 129 LOEGYENSRCYNEMALIRVQVS 150
DB 129 LYK--ENKREYKRVKACVQS 148

RESULT 22
US-08-795-303-9
Sequence 9, Application US/08795303
Patent No. 5948656
GENERAL INFORMATION:
APPLICANT: Anderson, Paul J.
TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, NW Suite 800
CITY: Washington
STATE: DC


```

; COUNTRY: USA
; ZIP: 20047
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795.303
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/318,947
; FILING DATE: 06-OCT-1994
; APPLICATION NUMBER: 08/133,530
; FILING DATE: 07-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,952
; REFERENCE/DOCKET NUMBER: A6462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)253-2920
; FAX: 6491103
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-795-303-9

Query Match 11.5%; Score 151; DB 2; Length 151;
Best Local Similarity 28.2%; Pred. No. 6.9e-09;
Matches 40; Conservative 30; Mismatches 58; Indels 14; Gaps 4;

QY 9 PEGIMVKTFEDRMDFSA:--ALIKGPTPTPYEDGLYLFDIQLPNLYPAVPPHFCYLSQCSGR 68
DB 21 PTGSSGAPDNNIMNNAVTFGHPDPTDCTFKLTIEETEEYPNKPPTVREVSQV---F 77
QY 69 LNPNDYNGKVCVSLGTTW:CKGTEHWTSKSSLLQVLISIOGLILVNEPYNEAGFDSRG 128
DB 78 HFNVTADGGICGLDI:-----GNKSPRYDVSAILISQSL--SDPNPNSPANSTAAQ 128
QY 129 LQEGYNSRCYNEMALIRVVS 150
DB 129 LYK--ENKRYEKVKACVQS 148

RESULT 23
US-08-767-942A-25
; Sequence 25, Application US/08767942A
; Patent No. 6068982
; GENERAL INFORMATION:
; APPLICANT: Rolfe, Mark
; APPLICANT: Chiu, M. Isabel
; APPLICANT: Berlin, Vivian
; APPLICANT: Damagone, Veronique
; APPLICANT: Draetta, Giulio
; APPLICANT: Guillaume, Cottare
; TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/767,942A
; FILING DATE: 17-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-029,04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-767-942A-25

Query Match 11.3%; Score 148; DB 3; Length 298;
Best Local Similarity 30.3%; Pred. No. 4.2e-08;
Matches 44; Conservative 28; Mismatches 55; Indels 18; Gaps 7;

QY 10 EGIWVKTFEDRMDFLS--ALIKGPTPTPYEDGLYLFDIQLPNLYPAVPPHFCYLSQCSGR 67
DB 88 EGERV-TLVDEGLYNWVAIFGPPNTYEGGYFKARKLPFDIDYPSPPAFRLTK---M 143
QY 68 LNPNDYNGKVCVSLGTTW:CKGTEHWTSKSSLLQVLISIOGLILVNEPYNE 120
DB 144 WHPNIYETDVCISILHPPYDDPQSGELPSEHWNPQNVKTIILSV--ISLLNEPNTSP 201
QY 121 AGFSDS---RGLQEGYNSRCYNEM 142
DB 202 ANVDASVMYRWKESKCKGKREYTDI 226

RESULT 24
US-09-177-165A-23
; Sequence 23, Application US/09177165A
; Patent No. 6426205
; GENERAL INFORMATION:
; APPLICANT: Iyers, Mike
; APPLICANT: Williams, Andrew
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
; TITLE OF INVENTION: DEPENDENT PROTEOLYSIS
; FILE REFERENCE: 11757.100U1
; CURRENT APPLICATION NUMBER: US/09/177,165A
; CURRENT FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/392,443
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/063,254
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 23
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-09-177-165A-23

Query Match 11.3%; Score 148; DB 4; Length 298;
Best Local Similarity 30.3%; Pred. No. 4.2e-08;
Matches 44; Conservative 28; Mismatches 55; Indels 18; Gaps 7;

QY 10 EGIWVKTFEDRMDFLS--ALIKGPTPTPYEDGLYLFDIQLPNLYPAVPPHFCYLSQCSGR 67
DB 88 EGERV-TLVDEGLYNWVAIFGPPNTYEGGYFKARKLPFDIDYPSPPAFRLTK---M 143
QY 68 LNPNDYNGKVCVSLGTTW:CKGTEHWTSKSSLLQVLISIOGLILVNEPYNE 120
DB 144 WHPNIYETDVCISILHPPYDDPQSGELPSEHWNPQNVKTIILSV--ISLLNEPNTSP 201

```

QY 12: AGEDSO---RGLQEGYNSRCYNEM 142
 DB 202 ANVDASVMYKWKFSKGKGRYTDI 226

RESULT 25
 US-09-058-368 2
 : Sequence 2, Application US/09058368
 : Patent No. 6277568

GENERAL INFORMATION:
 : APPLICANT: Lal, Preeti
 : APPLICANT: Hillman, Jennifer L.
 : APPLICANT: Guegler, Karl J.
 : APPLICANT: Cooley, Neil C.
 : APPLICANT: Baugh, Maria R.
 : APPLICANT: Azirizai, Valda
 : TITLE OF INVENTION: HUMAN URIQUILIN-CONJUGATING ENZYME
 : TITLE OF INVENTION: HOMOLGSS
 : NUMBER OF SEQUENCES: 6
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 374 Porter Drive
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304

COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSEQ for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/358,368
 : FILING DATE: BEKWITH
 : CLASSIFICATION:
 : PRIOR APPLICATION NUMBER:
 : FILING DATE:

ATTORNEY/AGENT INFORMATION:
 : NAME: Gertone, Michael C.
 : REGISTRATION NUMBER: 39,142
 : REFERENCE/DOCKET NUMBER: PF 0502 US
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 650-855-0555
 : TELEFAX: 650-845-4156
 : TELEX:

INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 282 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : LIBRARY: SIMTNOT13
 : CLONE: 1803903
 : US-09-358-368-2

Query Match 11.0%; Score 145.5; DB 4; Length 282;
 Best Local Similarity 32.0%; Pred. No. 7.3e-08;
 Matches 40; Conservative 25; Mismatches 45; Indels 15; Gaps 6;

QY 10 RGIWKTFEDRMOLFS--ALIKGPTRTYDGLYFDIQLPNITYPVPPHFCYLSQSGRL 67
 DB 26 RGFRI-TLVDESILYNNVAIFGLPNITLYEGGYKAKHKFPIDYPSYPTFRITK---M 81

QY 68 LNPNDYNGKVCVSLTGWTK-----GTERWTSKSLQLVLISIQGLILVNEP-YYNE 120
 DB 82 WHENTYNGDVCTSLHPVDDPSSE-FSERWNPNTQWRTILSV--ISLNEPTFSP 139

QY 12: AGFDS 125
 DB 140 ANVDA 144

RESULT 26
 US-09-358-580-12
 : Sequence 12, Application US/09358580
 : Patent No. 6107545
 : GENERAL INFORMATION:
 : APPLICANT: Mahajan, Pramod B.
 : TITLE OF INVENTION: Maize Rad6 Genes and Uses Thereof
 : FILE REFERENCE: 0883
 : CURRENT APPLICATION NUMBER: US/09/358,580
 : CURRENT FILING DATE: 1999-07-21
 : EARLIER APPLICATION NUMBER: 60/096,546
 : EARLIER FILING DATE: 1998-08-14
 : NUMBER OF SEQ ID NOS: 14
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 12
 : LENGTH: 152
 : TYPE: PRT
 : ORGANISM: Zea mays
 : US-09-358-580-12

Query Match 11.0%; Score 145; DB 3; Length 152;
 Best Local Similarity 28.6%; Pred. No. 3.2e-08;
 Matches 42; Conservative 27; Mismatches 58; Indels 20; Gaps 4;

QY 9 PEGIMVKTEDRMOLFSALIKGPTRTYDGLYFDIQLPNITYPVPPHFCYLSQSGRL 68
 DB 21 PAGISGAPQDNIMLWNAVTFGDDTGWGTEKLTLOPNEEYPNKPTVRFVSR---MF 77

QY 69 NPNYDNGKVCVSLTGWTKGIERWTSKSLQLVLISIQGLIL---VNEPYNEAGFDS 125
 DB 78 HPNIYADGSGICLDL-----QDQSPYDVAILTISQSLCLCPNPNSPANSAA--- 127

QY 126 DRGLQEGYNSRCYNEMALIRVQSWT 152
 DB 128 ----RMFSENKREYNKRYEVVEQSWT 150

RESULT 27
 US-08-533-298-2
 : Sequence 2, Application US/08533298
 : Patent No. 5851791
 : GENERAL INFORMATION:
 : APPLICANT: Vlerstra, Richard D
 : APPLICANT: Gosink, Mark M
 : TITLE OF INVENTION: Ubiquitin Conjugating Enzyme (E2) Fusion
 : TITLE OF INVENTION: Proteins
 : NUMBER OF SEQUENCES: 18
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Charles and Brady
 : STREET: 1 South Pinckney Street - Suite 600
 : CITY: Madison
 : STATE: WI
 : COUNTRY: USA
 : ZIP: 53703
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/533,298
 : FILING DATE: 25-SEP-1995
 : CLASSIFICATION: 536
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/070,157
 : FILING DATE: 28-MAY-1993
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Scay, Nicholas J
 : REGISTRATION NUMBER: 27,386
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 608-251-5000
 : TELEFAX: 608-251-9166

```

: INFORMATION FOR SEQ ID NO: 2:
:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 152 amino acids
:   TYPE: amino acid
:   TOPOLOGY: linear
:
: MOLECULE TYPE: protein
: US-08-533-298-2

```

```

Query Match      11.0%; Score 144.5; DR 2; Length 152;
Best local Similarity 26.4%; Pred. No. 3.6e-08;
Matches 42; Conservative 29; Mismatches 53; Indels 35; Gaps 5;

QY    5 PEGTMMVTFEDRMULFSALIKGPTREYDGLFLFDIQLPNITYPAVPPHCYLSOCSGRL 68
      || ..||..||..||..||..||..||..||..||..||..||..||..||..||
Db     21 FAGISGAQNIMINMNAIVFGPDITWOGGIKSLQFSDHYPKNTTVRFYSR---MF 77

QY    65 NNIIYNCKVCVSIMQTIGKTGRHWKSKSLLQVLITSTOGLIL---VNPEFYNEAGFDS 125
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     76 HPIINYAKSSCIDL-----QNCWPSPYDVAAILTQSLLCDPNPNSPANSER---- 126

QY    126 DRGTQEYNSRCYNEMALIRVVUSGMQIVKKRPVEVFQE 164
      ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     127 ---ARMYSE-----SKREYARKRVDRVVEQ 147

```

```

RESU: 26
US-38-905-689A-3
: Sequence 3, Application US/08965689A
: Patent No. 6015702
: GENERAL INFORMATION:
: APPLICANT: Lal, Proeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Corley, Nelli C.
: TITLE OF INVENTION: HUMAN JH1U1TIN CONJUGATING ENZYMES
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Inevite Pharmaceuticals, Inc.

```

```

Query Match 10.5%; Score 137.5; DB 3; Length 165;
Best Local Similarity 29.2%; Pred. No. 2.5e-07;
Matches 38; Conservative 27; Mismatches 52; Indels 13; Gaps
QY 4 LATSLEPGIWKTF-EDRMDFLSALIKIGTRTPYEDGLYLFDIQLNPYAPVPHFYCLS 62
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 16 LTLPPPGIVAGPMNEENFEWALINGMPDTCFEFGVFATLSPFLDYLPSPKRMFTC 75
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 63 CCSGRLPNLYONGKVCVSLIGT-----WICKGTERTKSSLLQVLIISTQGLILNWP 116
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 76 E--MHPNPYDPGRVCISILHAPCDPPGCVSYSSAEHSPVQSVKILISVVSMLA--EP 130
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 117 YNPEAGDSD 126
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 131 -NDESGANVD 139

```

```

: RESULT 29
: US-09-359-467-3
: ; Sequence 3, Application US/09359967
: ; Patent No. 6146624
: ; GENERAL INFORMATION:
: ; APPLICANT: Lal, Preeti
: ; APPLICANT: Hillman, Jennifer L.
: ; APPLICANT: Cozley, Neil C.
: ; TITLE OF INVENTION: HUMAN UBIQUITIN CONJUGATING ENZYMES
: ; NUMBER OF SEQUENCES: 6
: ; CORRESPONDENCE ADDRESS:
: ; ADDRESSEE: Incyte Pharmaceuticals, Inc.

```

Query Match 10.5%; Score 137.5; DB 4; Length 165;
 Rest local Similarity 29.2%; Prod. No. 2.5e-07;
 Matches 38; Conservative 27; Mismatches 52; Indels 13; Gaps 5;
 Qv 4 LAISLPEGIWVKT-EDRMOLFSAIKGTPPTPYEDGLYFLDIQLPNITYAPVPHFCYLS 62

```

: Patent NO. 6107545
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad6 Genes and Uses Thereof
: FILE REFERENCE: 0883
: CURRENT APPLICATION NUMBER: US/09/358,580
: CURRENT FILING DATE: 1999-07-21
: EARLIER APPLICATION NUMBER: 60/096,546
: EARLIER FILING DATE: 1998-08-14
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 148
: TYPE: PRT
: ORGANISM: Zea mays
US-09-358-580-2

Query Match          10.0%; Score 132; DB 3; Length 148;
Best Local Similarity 30.9%; Pred.No. 8.5e-07;
Matches 30; Conservative 20; Mismatches 35; Indels 12; Gaps

QY  22 DLF--SALIKGTRTPYEIGLYLFDIQLPNLYPAVPHFCYLSQCSGRLNPNIYDNGKVC 79
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  29 DMFHQWATIMGPPDPSYAGGVFLNIIHFPDPYFKPKVSKTKV---PFPNINSNGSIC 85
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  80 VSLIGTWIGKGTERTWTSKSSLLQVLISIQGLLIIVNEP 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  86 LDIL-----KQNSPALTISKVLLSICSLLDTRNP 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RSU:T 32
US-09-358-580-4
: Sequence 4, Application US/09358580
: Patent No. 6107545
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad6 Genes and Uses Thereof
: FILE REFERENCE: 0883
: CURRENT APPLICATION NUMBER: US/09/358,580
: CURRENT FILING DATE: 1999-07-21
: EARLIER APPLICATION NUMBER: 60/096,546
: EARLIER FILING DATE: 1998-08-14
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 148
: TYPE: PRT
: ORGANISM: Zea mays
US-09-358-580-4

Query Match          10.0%; Score 132; DB 3; Length 148;
Best Local Similarity 30.9%; Pred.No. 8.5e-07;
Matches 30; Conservative 20; Mismatches 35; Indels 12; Gaps

QY  22 DLF--SALIKGTRTPYEIGLYLFDIQLPNLYPAVPHFCYLSQCSGRLNPNIYDNGKVC 79
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  29 DMFHQWATIMGPPDPSYAGGVFLNIIHFPDPYFKPKVSKTKV---PFPNINSNGSIC 85
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  80 VSLIGTWIGKGTERTWTSKSSLLQVLISIQGLLIIVNEP 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  86 LDIL-----KQNSPALTISKVLLSICSLLDTRNP 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 33
US-09-358-580-6
: Sequence 6, Application US/09358580
: Patent No. 6107545
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad6 Genes and Uses Thereof
: FILE REFERENCE: 0883
: CURRENT APPLICATION NUMBER: US/09/358,580
: CURRENT FILING DATE: 1999-07-21

```

Db 28 DDMFHWQATINGPNDSPYGGVFELTIIHFTDYPKPKVAFTRI--YHPNINSNGSI 84

; GENERAL INFORMATION:
 ; APPLICANT: Rolle, Mark
 ; APPLICANT: Eckstein, Jens W.

```

1  APPLICANT:  Draetta, Giulio
2  APPLICANT:  Guillaume Cottarel
3  TITLE OF INVENTION:  Ubiquitin Conjugating Enzymes
4  NUMBER OF SEQUENCES:  21
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE:  IAHIVE & COCKFIELD
7  STREET:  60 State Street
8  CITY:  Boston
9  STATE:  MA
10 COUNTRY:  USA
11 ZIP:  02109
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  ASCII(Ext)
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER:  US/08/486,663A
19 FILING DATE:
20 CLASSIFICATION:
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER:  US
23 FILING DATE:
24 ATTORNEY/AGENT INFORMATION:
25 NAME:  Vincent, Matthew P.
26 REGISTRATION NUMBER:  36,739
27 REFERENCE/DOCKET NUMBER:  M12-023CP2
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE:  (617) 227-7400
30 TELEFAX:  (617) 227-5941
31 INFORMATION FOR SEQ ID NO:  2:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH:  147 amino acids
34 TYPE:  amine acid
35 TOPOLOGY:  linear
36 MOLECULE TYPE:  protein
37 US 08-486 663A-2

```

```

Query Match          9.9%:  Score 130;  DB 2;  Length 147;
Best Local Similarity 28.6%:  Pred. No. 1.4e-06;
Matches 28;  Conservative 20;  Mismatches 40;  Indels 10;  Gaps 2;

QY  19 DRMDLPSALKGPTIRTPEDGILYFDIQLPNIVYAVPPHPCYLSOCGRINPNLYDNGKV 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  28 DDMFHQAATIMGNDSPYQGGVFFLTTHFTDYPPKPKVAFTRI---YHPNINSNGSI 84
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY  79 CVSLIGTWICKGTERWTSKSLLOVLISIQGLIIVNEP 116
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  85 CLDIL-----RSQSPALITISKVLLISICSLCDPNP 115
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 37
US-08-247-904B-2
: Sequence 2, Application US/08247904B
: Patent No. 5981699
: GENERAL INFORMATION:
: APPLICANT:  Rolfe, Mark
: APPLICANT:  Eckstein, Jens W.
: APPLICANT:  Draetta, Giulio
: TITLE OF INVENTION:  Human Ubiquitin Conjugating Enzyme
: NUMBER OF SEQUENCES:  17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE:  Foley, Hoag & Eliot
: STREET:  One Post Office Square
: CITY:  Boston
: STATE:  MA
: COUNTRY:  USA
: ZIP:  02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE:  Floppy disk
: COMPUTER:  IBM PC compatible
: OPERATING SYSTEM:  PC-DOS/MS-DOS
: SOFTWARE:  ASCII(text)

```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER:  US/08/247,904B
: FILING DATE:  23-MAY-1994
: CLASSIFICATION:  530
: ATTORNEY/AGENT INFORMATION:
: NAME:  Vincent, Matthew P.
: REGISTRATION NUMBER:  36,709
: REFERENCE/DOCKET NUMBER:  M1V-029,01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE:  (617) 832-1000
: TELEFAX:  (617) 832-7000
: INFORMATION FOR SEQ ID NO:  2:
: SEQUENCE CHARACTERISTICS:
: LENGTH:  147 amino acids
: TYPE:  amino acid
: TOPOLOGY:  linear
: MOLECULE TYPE:  protein
: US-08-247-904B-2

Query Match          9.9%:  Score 130;  DB 2;  Length 147;
Best Local Similarity 28.6%:  Pred. No. 1.4e-06;
Matches 28;  Conservative 20;  Mismatches 40;  Indels 10;  Gaps 2;

QY  19 DRMDLPSALKGPTIRTPEDGILYFDIQLPNIVYAVPPHPCYLSOCGRINPNLYDNGKV 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  28 DDMFHQAATIMGNDSPYQGGVFFLTTHFTDYPPKPKVAFTRI---YHPNINSNGSI 84
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY  79 CVSLIGTWICKGTERWTSKSLLOVLISIQGLIIVNEP 116
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  85 CLDIL-----RSQSPALITISKVLLISICSLCDPNP 115
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 38
US-08-895-601-8
: Sequence 8, Application US/08895601
: Patent No. 6060262
: GENERAL INFORMATION:
: APPLICANT:  Beer-Romero, Peggy
: APPLICANT:  Strack, Peter J.
: APPLICANT:  Glass, Susan J.
: APPLICANT:  Rolfe, Mark
: TITLE OF INVENTION:  REGULATION OF KAPPA B (KB) DEGRADATION,
: TITLE OF INVENTION:  AND METHODS AND REAGENTS RELATED THERETO
: NUMBER OF SEQUENCES:  16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE:  FOLEY, HOAG & ELIOT LLP
: STREET:  One Post Office Square
: CITY:  Boston
: STATE:  MA
: COUNTRY:  USA
: ZIP:  02109-2170
: COMPUTER READABLE FORM:
: MEDIUM TYPE:  Floppy disk
: COMPUTER:  IBM PC compatible
: OPERATING SYSTEM:  PC-DOS/MS-DOS
: SOFTWARE:  PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER:  US/08/895,601
: FILING DATE:  16-JUL-1997
: CLASSIFICATION:  435
: ATTORNEY/AGENT INFORMATION:
: NAME:  Vincent, Matthew P.
: REGISTRATION NUMBER:  36,709
: REFERENCE/DOCKET NUMBER:  M1V-096,01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE:  617-832-1000
: TELEFAX:  617-832-7000
: INFORMATION FOR SEQ ID NO:  8:
: SEQUENCE CHARACTERISTICS:
: LENGTH:  147 amino acids
: TYPE:  amino acid
: TOPOLOGY:  linear
: MOLECULE TYPE:  protein

```

98-023-60-50

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:29:19 ; Search time 39.7059 seconds
(without alignments)
1297.334 Million cell updates/sec

Title: US-09-930-026-1

Perfect score: 1315

Sequence: 1 MALLATSLPGRHWKTPEDR.....PGASGDSSEGAOGLAFS 250

Scoring table: H10SUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 135 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mnc.*
- 8: sp_orquaele.*
- 9: sp_plage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1301	98.9	466	4	Q9H9B2 homo sapien
2	1299	98.8	742	4	Q9H6H6
3	1295	98.8	1313	4	Q9C0C9
4	1286	97.8	466	4	Q8TBN1
5	1202	91.4	311	4	Q9H7F4
6	87	66.2	255	4	Q9H5W1
7	551.5	41.9	1398	5	Q9XZ32
8	474	36.0	1102	10	Q9ZVX1
9	451.5	34.3	267	10	Q9AWU5
10	427.5	32.5	907	10	Q9V110
11	425.5	32.4	540	10	Q9LGG0
12	423.5	32.2	543	10	Q9LNM8
13	405.5	30.8	273	10	Q9M7X2
14	405.5	30.8	608	10	Q9L0U5
15	404.5	30.8	359	10	Q9LGY0
16	391	29.7	182	10	P93012

17	376.5	28.6	323	10	Q9C917
18	374.5	28.5	244	10	Q9C918
19	372.5	28.3	4504	5	Q9VH01
20	372	28.3	493	10	Q8S193
21	372	28.3	4845	11	Q88738
22	370	28.1	1867	4	Q9H8B7
23	248	18.9	246	4	Q9H832
24	248	18.9	246	11	Q8R214
25	244	18.6	236	4	Q9H893
26	242	18.4	232	11	Q9DAU4
27	169	12.9	222	4	Q9BTC1
28	168	12.8	153	10	Q9XHP3
29	166	12.6	153	10	Q9A497
30	164	12.5	153	10	Q9E248
31	164	12.5	197	4	Q9NPD8
32	163	12.4	162	5	Q8SR17
33	162	12.3	166	3	Q9C2A5
34	160	12.2	152	4	Q96FX4
35	160	12.2	152	11	Q9Z255
36	160	12.2	152	13	Q9W6F3
37	160	12.2	204	11	Q8CQ37
38	158.5	12.1	151	3	P78717
39	158.5	12.1	190	5	Q9VX25
40	157.5	12.0	329	5	Q9SXN7
41	157	11.9	152	11	Q9D0Z6
42	156	11.9	126	6	Q9NIX7
43	156	11.9	153	10	Q8W011
44	155.5	11.8	238	11	Q8VDE5
45	154	11.7	152	10	Q9APV0
46	154	11.7	152	10	Q9AVN9
47	153.5	11.7	223	11	Q9Z1J4
48	153.5	11.7	560	10	Q8W0K6
49	153	11.6	151	3	Q9G0P5
50	152	11.6	126	6	Q9N1X6
51	151	11.5	163	10	Q9ZVA6
52	150	11.4	154	10	Q9C8X7
53	149.5	11.4	151	5	Q964Q5
54	149.5	11.4	238	4	Q9NX64
55	149	11.3	124	6	Q9N1X5
56	149	11.3	170	5	Q9U1Q1
57	148.5	11.3	235	5	Q76542
58	148	11.3	152	10	Q9M4R0
59	148	11.3	152	10	Q8SBC1
60	147	11.2	124	6	Q9SL52
61	147	11.2	355	10	Q9AUL4
62	145.5	11.1	199	5	Q9T269
63	145	11.0	341	5	Q9VUH4
64	144	11.0	148	3	Q13685
65	143	10.9	354	5	Q917T6
66	142.5	10.8	251	10	Q9FF66
67	141.5	10.8	217	3	Q74810
68	140.5	10.7	172	5	Q8S554
69	139.5	10.6	148	10	Q9SPH9
70	139.5	10.6	199	5	Q8T0Z6
71	137.5	10.5	200	11	Q54806
72	137.5	10.5	192	10	Q9F161
73	137	10.4	160	5	Q77397
74	137	10.4	160	5	Q77397
75	136.5	10.4	148	11	Q9CX59
76	136	10.3	144	10	Q9AW53
77	136	10.3	199	5	P91633
78	135.5	10.3	148	5	Q9NED4
79	135.5	10.3	167	5	Q9VXE8
80	135.5	10.3	168	5	Q8SVG3
81	135.5	10.3	200	5	Q9N9Z5
82	135	10.3	152	11	Q9DAJ6
83	134	10.2	178	10	Q9FPH9
84	134	10.2	186	5	Q9U5Q3
85	134	10.2	195	10	Q8S1Y5
86	133.5	10.2	146	10	Q8SH72
87	133.5	10.2	148	10	Q94F47
88	133	10.1	147	10	Q8S920
89	133	10.1	148	10	Q42897

Q9C917	arabidopsis
Q9C918	arabidopsis
Q9VH01	drosophila
Q8S193	oryza sativ
Q88738	mus musculus
Q9H8B7	homo sapien
Q9H832	homo sapien
Q8R214	mus musculus
Q9H893	homo sapien
Q9BTC1	homo sapien
Q9XHP3	catharanthu
Q9A497	arabidopsis
Q9E248	arabidopsis
Q9NPD8	homo sapien
Q8SR17	encephalito
Q9C2A5	neurospora
Q96FX4	homo sapien
Q9Z255	mus musculus
Q9W6F3	gallus gall
Q9CQ37	mus musculus
P78717	nectria hae
Q9VX25	drosophila
Q9SXN7	caenorhabdi
Q9D0Z6	mus musculus
Q9NIX7	sus scrofa
Q8W011	oryza sativ
Q8VDE5	mus musculus
Q9APV0	nicotiana t
Q9AVN9	nicotiana t
Q9Z1J4	mus musculus
Q8W0K6	oryza sativ
Q9G0P5	emericea
Q9N1X6	equus cabal
Q9ZVA6	arabidopsis
Q9C8X7	arabidopsis
Q964Q5	trypanosoma
Q9NX64	homo sapien
Q9N1X5	bos taurus
Q9U1Q1	caenorhabdi
Q76542	dictyosteli
Q9M4R0	avicennia m
Q8SBC1	oryza sativ
Q9SL52	bos taurus
Q9AUL4	oryza sativ
Q9T269	caenorhabdi
Q9VUH4	drosophila
Q13685	scizosacch
Q917T6	drosophila
Q9FF66	arabidopsis
Q74810	schizosacch
Q8S554	encephalito
Q9SPH9	mesembryant
Q8T0Z6	bombyx mori
Q54806	mus musculus
Q9F161	arabidopsis
Q9C9X9	rattus norv
Q77397	plasmodium
Q9CX59	mus musculus
Q9AW53	guillardia
P91633	drosophila
Q9NED4	leistmania
Q9VXE8	drosophila
Q8SVG3	drosophila
Q9N9Z5	drosophila
Q9DAJ6	mus musculus
Q9FPH9	arabidopsis
Q9U5Q3	entodinium
Q8S1Y5	oryza sativ
Q8SH72	arabidopsis
Q94F47	arabidopsis
Q8S920	oryza sativ
Q42897	lycopersico


```

90 133 10.1 148 10 Q5SLE4
91 133 10.1 374 10 Q94P88
92 132 10.0 148 10 Q48555
93 131.5 10.0 140 10 Q8VWV2
94 131.5 10.0 185 5 Q95X00
95 131.5 10.0 491 5 Q9V165
96 131 10.0 148 10 Q43821
97 131 10.0 257 3 Q8X0N3
98 130.5 9.9 146 10 Q94B2
99 130 9.9 147 4 Q9V2X8
100 130 9.9 147 11 Q907F5
101 130 9.9 147 11 Q9D1S1
102 129 9.9 148 5 Q9GNC1
103 129 9.8 118 11 Q9DB08
104 128 9.7 148 10 Q9FKT3
105 127.5 9.7 148 10 Q8S9J9
106 127 9.7 148 4 Q96RP6
107 127 9.7 148 11 Q9D6V0
108 127 9.7 167 3 Q9V818
109 127 9.7 194 10 Q43780
110 126.5 9.6 270 5 Q960Q5
111 125.5 9.5 129 10 Q93722
112 125.5 9.5 147 3 Q74196
113 125.5 9.5 148 10 Q9Z128
114 125.5 9.5 151 5 Q9V1S5
115 125.5 9.5 231 5 Q9NKC1
116 125 9.5 147 3 Q9JVR2
117 125 9.5 147 10 Q9RXA0
118 125 9.5 148 10 Q8XQ00
119 124.5 9.5 177 10 Q48838
120 124.5 9.5 184 4 Q9BQ25
121 124 9.4 409 10 Q9LY54
122 124 9.4 409 10 Q941H6
123 123.5 9.4 148 5 Q94490
124 123 9.4 159 5 Q62622
125 123 9.4 168 5 Q9VFR8
126 122.5 9.3 179 11 Q9D1C1
127 122 9.3 147 10 Q957H6
128 120.5 9.2 160 10 Q9W2X7
129 120.5 9.2 166 10 Q42541
130 120.5 9.2 178 5 Q9V7Y6
131 120 9.1 166 5 Q95017
132 120 9.1 160 10 Q42551
133 119.5 9.1 182 10 Q9LJD7
134 119 9.0 148 10 Q42973
135 118.5 9.0 126 10 Q24240

```

ALIGNMENTS

```

RESULT 1
Q9H9B2 PRELIMINARY; PRJ: 466 AA.
AC Q9H9B2;
DI 01-MAR-2001 (TRENBLrel. 16, Created)
DI 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DI 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE CDNA FLJ12878 fis, clone NR2RP2003840, weakly similar to hypothetical
DE 48.1 kDa protein B0403.2 in chromosome X.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Iogura S., Komai F., Hara K., Takuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahara K., Mascho Y., Oshima A.;
RT *NEDO human cDNA sequencing project.*;
RL Submitted (AUG-2000) to the ENBL/GenBank/DBJ databases.

```

```

DR EMBL: AK022940; BAB14320.1; --
DR HSSP: Q02159; 2UCZ.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBCC; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
SQ SEQUENCE 466 AA; 51544 MW; 84CEDD62B7843AF2 CRC64;

Query Match 98.9%; Score 1301; DB 4; Length 466;
Best Local Similarity 99.6%; Pred. No. 5, 1e-110;
Matches 247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MALLATSLPEGIMVTFEDRMDLFSALIKGTRTPRYEDGLYLFDIQLPNTIYPVPPHFCY 60
|||||
DB 136 MALLATSLPEGIMVTFEDRMDLFSALIKGTRTPRYEDGLYLFDIQLPNTIYPVPPHFCY 195

QY 61 LSQCSGRLNPNDYNGKVCVSLTGTWIGKTERWTSKSSLLQVLISIQGLILVNEPYNE 120
|||||
DB 196 LSQCSGRLNPNDYNGKVCVSLTGTWIGKTERWTSKSSLLQVLISIQGLILVNEPYNE 255

QY 121 AGFDSDRGLQEGYNSRCYNEMALIRVQSMOTLVRRPPEVEQRIKHQFSTIGCWLVRK 180
|||||
DB 256 AGFDSDRGLQEGYNSRCYNEMALIRVQSMOTLVRRPPEVEQRIKHQFSTIGCWLVRK 315

QY 181 IESWLETHALLEKAALPNGVPPKASSPEPPAVAEISDSGGQEPEDGGPAPGEASQSDS 240
|||||
DB 316 IESWLETHALLEKAALPNGVPPKASSPEPPAVAEISDSGGQEPEDGGPAPGEASQSDS 375

QY 241 EGGAGGLA 248
|||||
DB 376 EGGAGGLA 383

RESULT 2
Q9H6E6 PRELIMINARY; PRJ: 742 AA.
AC Q9H6E6;
DI 01-MAR-2001 (TRENBLrel. 16, Created)
DI 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DI 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE CDNA: FLJ22346 fis, clone HRC06158.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiiji T., Kobatake N., Inaaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT *NEDO human cDNA sequencing project.*;
RL Submitted (AUG-2000) to the ENBL/GenBank/DBJ databases.
DR EMBL: AK025999; BAB15313.1; --
DR HSSP: Q02159; 2UCZ.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBCC; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
SQ SEQUENCE 742 AA; 81876 MW; DDAEDB3RE66567C CRC64;

Query Match 98.8%; Score 1299; DB 4; Length 742;
Best Local Similarity 99.6%; Pred. No. 1, 4e-109;
Matches 247; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MALLATSLPEGIMVTFEDRMDLFSALIKGTRTPRYEDGLYLFDIQLPNTIYPVPPHFCY 60
|||||
DB 412 MALLATSLPEGIMVTFEDRMDLFSALIKGTRTPRYEDGLYLFDIQLPNTIYPVPPHFCY 471

QY 61 LSQCSGRLNPNDYNGKVCVSLTGTWIGKTERWTSKSSLLQVLISIQGLILVNEPYNE 120
|||||
DB 472 LSQCSGRLNPNDYNGKVCVSLTGTWIGKTERWTSKSSLLQVLISIQGLILVNEPYNE 531

```

QY 121 AGFSDRGLOQYENSKYCNEMALIRVQSMTOIVRRPPEVFEQEIHQHSTGGRRLLVNR 180
 |||||
 DB 532 AGFSDRGLOQYENSKYCNEMALIRVQSMTOIVRRPPEVFEQEIHQHSTGGRRLLVNR 591
 |||||
 QY 181 IESWLETHALLEKKAQALPNCVPRKASSPPPPAVAEISDSGQEPDGGPAPGEASQGS 240
 |||||
 DB 592 IESWLETHALLEKKAQALPNCVPRKASSPPPPAVAEISDSGQEPDGGPAPGEASQGS 651
 |||||
 QY 241 EGGAQSLA 248
 |||||
 DB 652 EGGAQSLA 659
 |||||
 RESULT 3
 ID Q9C0C9 PRELIMINARY: PRT: 1313 AA.
 AC Q9C0C9;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE KIAA1734 protein (Fragment).
 GN KIAA1734.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-21082932; PubMed-11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 KT "Prediction of the coding sequences of unidentified human genes. XIX.
 KT The complete sequences of 100 new cDNA clones from brain which code
 KT for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -1- M-SCHELANOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC TIOLESTER FORMATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DB EMBL: ARC51521; BAB21825.1;
 DB HSSR: Q22159; 20CZ
 DB InterPro: IPR000609; Ubq_conjugat.
 DB Pfam: PF00779; Ubq_con: 1.
 DB Prodom: PD000461; Ubq_conjugat: 1.
 DB SMART: SM00212; UBQC: 1.
 DB PROSITE: PS50127; UBIQUITIN_CONJUGANT_2: 1.
 KW Ligase; Ubiquitin conjugation.
 FI NON_TER
 SQ SEQUENCE 1313 AA; 143476 MW; 5A129203D878797C CRC64;

Query Match: 98.8%; Score 1299; DB 4; Length 1313;
 Best Local Similarity: 99.6%; Pred. No. 3, 10-109;
 Matches 247; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MALLATSLEPIGIMVKTEDMDLFSALIKGPTPTPYEDGLYFDIQLPNIYPVPPHFCY 60
 |||||
 DB 983 MALLATSLEPIGIMVKTEDMDLFSALIKGPTPTPYEDGLYFDIQLPNIYPVPPHFCY 1042
 |||||
 QY 61 LSQCSGRNPNDYNGKVCVSLIGTWIGKTERWTSKSSLIQVLIISQGLIILVNEPYNE 120
 |||||
 DB 1043 LSQCSGRNPNDYNGKVCVSLIGTWIGKTERWTSKSSLIQVLIISQGLIILVNEPYNE 1102
 |||||
 QY 121 AGFSDRGLOQYENSKYCNEMALIRVQSMTOIVRRPPEVFEQEIHQHSTGGRRLLVNR 180
 |||||
 DB 1103 AGFSDRGLOQYENSKYCNEMALIRVQSMTOIVRRPPEVFEQEIHQHSTGGRRLLVNR 1162
 |||||
 QY 181 IESWLETHALLEKKAQALPNCVPRKASSPPPPAVAEISDSGQEPDGGPAPGEASQGS 240
 |||||
 DB 1163 IESWLETHALLEKKAQALPNCVPRKASSPPPPAVAEISDSGQEPDGGPAPGEASQGS 1222
 |||||

QY 241 EGGAQSLA 248
 |||||
 DB 1223 EGGAQSLA 1230
 |||||
 RESULT 4
 ID Q8IBN1 PRELIMINARY: PRT: 466 AA.
 AC Q8IBN1;
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Likely ortholog of mouse ubiquitin-conjugating enzyme E2-230k.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-BRAIN;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC022237; AAH22237.1;
 SQ SEQUENCE 466 AA; 51477 MW; DE87F936AC995EC43 CRC64;
 Query Match: 97.8%; Score 1286; DB 4; Length 466;
 Best Local Similarity: 99.2%; Pred. No. 1, 2e-108;
 Matches 246; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MALLATSLEPIGIMVKTEDMDLFSALIKGPTPTPYEDGLYFDIQLPNIYPVPPHFCY 60
 |||||
 DB 136 MALLATSLEPIGIMVKTEDMDLFSALIKGPTPTPYEDGLYFDIQLPNIYPVPPHFCY 195
 |||||
 QY 61 LSQCSGRNPNDYNGKVCVSLIGTWIGKTERWTSKSSLIQVLIISQGLIILVNEPYNE 120
 |||||
 DB 196 LSQCSGRNPNDYNGKVCVSLIGTWIGKTERWTSKSSLIQVLIISQGLIILVNEPYNE 255
 |||||
 QY 121 AGFSDRGLOQYENSKYCNEMALIRVQSMTOIVRRPPEVFEQEIHQHSTGGRRLLVNR 180
 |||||
 DB 256 AGFSDRGLOQYENSKYCNEMALIRVQSMTOIVRRPPEVFEQEIHQHSTGGRRLLVNR 315
 |||||
 QY 181 IESWLETHALLEKKAQALPNCVPRKASSPPPPAVAEISDSGQEPDGGPAPGEASQGS 240
 |||||
 DB 316 IESWLETHALLEKKAQALPNCVPRKASSPPPPAVAEISDSGQEPDGGPAPGEASQGS 375
 |||||
 QY 241 EGGAQSLA 248
 |||||
 DB 376 EGGAQSLA 383
 |||||
 RESULT 5
 ID Q9H7E4 PRELIMINARY: PRT: 311 AA.
 AC Q9H7E4;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE CUNA: FLJ21004 fls, clone CAE03743.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 KT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK024657; BARI4948.1;
 DR HSSP: Q02159; 2UCZ.
 DB InterPro: IPR000608; Ubq_conjugat.

01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 BCDNA: LD22087 OR CG10254
 BCDNA: LD22087 OR CG10254
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta.
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A.
 STRAIN-BERKELEY:
 MEDLINE=2019606; PubMed=10731132;
 Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Ananadis P.G., Scher S.F., Li P.W., Hoskins R.A., Gallo R.P.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton C.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews-Plannkovich C., Baldwin D.,
 Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman H.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattler P.,
 Burtis K.M., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.C., Cawley S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunt P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Fortiera S., Fleischmann W.,
 Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 Hoskins D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Romington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 The genome sequence of Drosophila melanogaster.;
 Science 287:2185-2195(2000).
 [2]
 SEQUENCE FROM N.A.
 Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 Agbayani A., Arcina T.F., Baxter E., Blazaj R.G., Butenhoff C.,
 Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 Gallo R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
 Houston K.A., Hummasti S.K., Kim E., Li P., Moshrefi M., Pacle J.M.,
 Park S., Sequeira A., Sethi H., Snir E., Svrtkars R.R., Weinburg T.,
 Celnik S.E.;
 Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 PROTEINS (BY SIMILARITY).
 -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP -
 DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
 -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 THIOLESTER FORMATION (BY SIMILARITY).
 -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 EMBL: AF003744; AAF56162.1;
 EMBL: AF132190; AAD34778.1;
 HSP: P50550; 109A
 FyBase: FBqn0027512; BCDNA:LD22087.

01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 01-NAR-2002 (TrEMBLrel. 20, Last annotation update)
 BCDNA: LD22087 OR CG10254
 BCDNA: LD22087 OR CG10254
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta.
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-BERKELEY:
 RX MEDLINE=2019606; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananadis P.G., Scher S.F., Li P.W., Hoskins R.A., Gallo R.P.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton C.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman H.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P., Downes M., Dugan-Rocha S., Dunkov B.C., Dunt P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fortiera S., Fleischmann W.,
 RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck C.,
 RA Hoskins D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Romington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 The genome sequence of Drosophila melanogaster.;
 Science 287:2185-2195(2000).
 [2]
 RN SEQUENCE FROM N.A.
 RP Ruben G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 RA Agbayani A., Arcina T.F., Baxter E., Blazaj R.G., Butenhoff C.,
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 RA Gallo R., George R.A., Harris R.A., Hoskins R.A., Evans-Holm M.,
 RA Houston K.A., Hummasti S.K., Kim E., Li P., Moshrefi M., Pacle J.M.,
 RA Park S., Sequeira A., Sethi H., Snir E., Svrtkars R.R., Weinburg T.,
 RA Celnik S.E.;
 RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RL -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP -
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
 CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AF003744; AAF56162.1; -;
 DR EMBL: AF132190; AAD34778.1; -;
 DR HSP: P50550; 109A
 DR FyBase: FBqn0027512; BCDNA:LD22087.

DR InterPro: IPR000608: UBQ_conjugat.
 DR Pfam: PF00179; UQ_con: 1.
 DR ProDom: PD000461: UBQ_conjugat: 1.
 DR SMART: SM00212; UBCc: 1.
 DR PROSITE: PS00639: THIOL_PROTEASE_HIS; UNKNOWN_1.
 DR PROSITE: PS0127: UBIQUITIN_CONJUGAT_2; 1.
 KW Hypothetical protein; Ubiquitin conjugation.
 SQ SEQUENCE: 1398 AA; 15645 MW; RE50C553AACBICP CRC64;

Query Match 41.9%; Score 551.5; DB 5; Length 1398;
 Best Local Similarity 49.4%; Pred. No. 3.4e-41;
 Matches 104; Conservative 40; Mismatches 64; Indels 7; Caps 2;

QY 3 LIATSLPGIMVKTFEDRMDFLSALIKGPTPTPYDGLYFDIQLPNITYPAVPPHFCYLS 62
 Db 1132 MIKSLNGVYVAYEDRMDFLSVWVGPKPTPYNALFFEDQGRDYPKSPVPCVHS 1191
 QY 63 QCSGRNPNLYDNGKVCVSLGNTGKGTERTSKSSLLQVLISQGLLVNPPYNEAG 122
 Db 1192 VCTDRINPNLYDNGKVCVSLGNTGKGTERTSKSSLLQVLISQGLLVNPPYNEAG 1251
 QY 123 FDSRGTLQEGYENSCYNEALIRVQSKTQLYRRPPEVEQEIHQHSTGGWHLVNR 182
 Db 1232 YEKURGTLQENSRVYNEALIRVQSKTQLYRRPPEVEQEIHQHSTGGWHLVNR 182
 QY 183 SWLETHAL-----LKAQALDNGVPKASSSPPP 211
 Db 1312 AKSE-YSLAQRONITKIDMEVEXKASCELPPEP 1345

RESULT 8

Q9ZVX1 ID Q9ZVX1 PRELIMINARY; PRI; 1102 AA.
 AC Q9ZVX1
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative ubiquitin-conjugating enzyme.
 GN A72016920.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 QX NCBI_taxid 3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUDB-A;
 RX MEDLINE-20083487; PubMed-10617197;
 RA Lin X., Kaul S., Mounslay S.D., Shea T.P., Henito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Howman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Hue., C.R., Ketchum K.A., Lee J.J., Koning C.M., Koo H., Moffat K.S.,
 RA Cronin J.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy I.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RI "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana";
 RT Nature 402:761-768(1999).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-CV. COLUMBIA;
 RC Lin X.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOL-ESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AC005167; AAC64223.1; -;
 DR HSSP: Q02159; 2UC4.
 DR InterPro: IPR000169; Shprot_acsite.

DR InterPro: IPR000608: UBQ_conjugat.
 DR Pfam: PF00179; UQ_con: 1.
 DR ProDom: PD000461: UBQ_conjugat: 1.
 DR SMART: SM00212; UBCc: 1.
 DR PROSITE: PS00639: THIOL_PROTEASE_HIS; UNKNOWN_1.
 DR PROSITE: PS0127: UBIQUITIN_CONJUGAT_2; 1.
 KW Ubiquitin conjugation.
 SQ SEQUENCE: 1102 AA; 122182 MW; DA2C11385355E8D9 CRC64;

Query Match 36.0%; Score 474; DB 10; Length 1102;
 Best Local Similarity 47.1%; Pred. No. 2.8e-34;
 Matches 89; Conservative 40; Mismatches 58; Indels 2; Caps 2;

QY 3 LIATSLPGIMVKTFEDRMDFLSALIKGPTPTPYDGLYFDIQLPNITYPAVPPHFCYLS 62
 Db 861 ILQNLPDGIQVAYEDRMDFLSALIKGPTPTPYDGLYFDIQLPNITYPAVPPHFCYLS 62
 QY 63 QCSGRNPNLYDNGKVCVSLGNTGKGTERTSKSSLLQVLISQGLLVNPPYNEA 121
 Db 920 SCGRNPNLYDNGKVCVSLGNTGKGTERTSKSSLLQVLISQGLLVNPPYNEA 121
 QY 122 FDSRGTLQEGYENSCYNEALIRVQSKTQLYRRPPEVEQEIHQHSTGGWHLVNR 181
 Db 980 GYDKQVGTARGENKSLCYNEALIRVQSKTQLYRRPPEVEQEIHQHSTGGWHLVNR 181
 QY 182 ESWLETHAL 190
 Db 1040 DAYMKGYLI 1048

RESULT 9

Q9AWU5 ID Q9AWU5 PRELIMINARY; PRT; 1067 AA.
 AC Q9AWU5
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE P0044F08.17 protein (P0037C04.30 protein).
 GN P0044F08.17 OR P0037C04.30.
 OS Oryza sativa (rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_taxid 4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki I., Matsumoto T., Yamamoto K.;
 RI "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0044F08.*";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RI "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0037C04.*";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOL-ESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AP002909; BAB21187.1; -;
 DR EMBL: AP003233; BAB55541.1; -;
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00179; UQ_con: 1.
 DR ProDom: PD000461; UBQ_conjugat: 1.
 DR SMART: SM00212; UBCc: 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.

KW Lique: Ubiquitin conjugation.
 SQ SEQUENCE 1067 AA: 117523 MW: 5A03E28A38102BBS CRC64;
 Query Match 34.38; Score 451.5; DH 10; Length 1067;
 Best Local Similarity 42.48; Pred. No. 3c-32;
 Matches 98; Conservative 39; Mismatches 79; Indels 15; Gaps

QY	3	LLAISPEGIWKTEDRMDFSALINGTRKPYEDGZYLFDIQLPNIPVAPVPEFCYLS 62
DB	821	ILEKNDPYIYVRVEDRMIDITRAVLIGASG:PYDQG:FFDFHLPFPQVPPSAYHS 880
QY	63	QCSG-RLNPMLYDNKGVYSLGTWICKGTERW-TSKSSLLQVZLSUGLILVNEPYNE 120
DB	681	--GSLRWPLXYDGVKGLSLENTWGRNEVMDPSSSILQVLVSLOGVLNEKPYNE 938
QY	121	AGFDSNCLQGGYNSRCYNEMALIRVQSGMTQLVRRPPEVEQEIRHOHSTGGWRLVNR 180
DB	939	AGYEKQVTVGEGKNALPNENIYLSLSKMYLITRRPPEHFDFAKSHKRGKYLKA 998
QY	181	IESMLFTHAL-----LEKAQALNPYKPKASSPEPPAVABELSDG 220
DB	999	CEAVILQNGVCTLIDIACTERSKQPCSGFKIALAKIMPRITALKDAG 1049

RESULT 10
 QRYVIC PRELIMINARY; PRI: 907 AA.
 ID Q8VY13;
 AC Q8VY13;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative ubiquitin-conjugating enzyme E2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 GN Arabidopsis
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chau M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Full length cDNA of gene At2g33770 (GI:1528164).
 RI Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RT *Arabidopsis open Reading Frame (ORF) clones.
 RI Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Hanh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,

[illegible]

```

RESULT 14
Q9LU05 PRELIMINARY; PRT; 609 AA.
AC Q9LU05;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Similarity to ubiquitin-conjugating enzyme (A13G15355/MJK13_1).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Caryophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsais.
OX NCBI_taxonomy/3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MED:119-20277480; PubMed=10819329;
RA Nakamura Y.;
RI "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RI features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RI clones."
RI DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Banth J., Bowser J.

```

```

OC Ehrhartoidae; Oryzaceae; Oryza.
ON NCBI_TaxID=4530;
OX [1]
RY SEQUENCE FROM N.A.
RP STRAIN-CV. NIPPONBARE;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare(CA3) genomic DNA, chromosome 1, PAC
RI clone: P070F03.*";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE ~ AMP +
CC DIPOSPHATE + PROTEIN N-UBIQUITYLYSINE.
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR HMBL: AP002481; BAA96594.1;
DR HSSP: P50550; I09B.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UQ_conj_1.
DR ProDom: PD000461; UBQ_conjugat.1.
DR SMART: SM00212; UBCC_1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW ligase; Ubiquitin conjugation.
SQ SEQUENCE 359 AA; 40029 MW; EH61A0D6D0F62F0 CRC64;

Query Match 30.8%; Score 404.5; DB 10; Length 359;
Best Local Similarity 40.2%; Pred.No. 1.3e-28;
Matches 88; Conservative 34; Mismatches 70; Indels 27; Gaps

QY 10 EGIMVKTFIFRMDFSAIIKGTRTPYEDGLYLFDIQLPNITYPAPVPHECYLSOCG-RL 68
Db 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 ESIVYRAYEDRIDLLPAALVGAGTYPHGELFFDVDPPEYFGSPKVVHS--GGRL 173
QY 69 NPNIYDKGVKCVSLGTWIGKGERW-TSKSSLQVLITSQGILLVNEPYNEACFPDSR 127
Db 17 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 NPNIYESGVKVCISLNTWMGTGCEKCKSKSTILOVLSIQGLVNLNDKIYENPGNKNSA 233
QY 128 GLOEGYENSRCYNEMALIRVVQSMTOLVARPPPEVEEIQIHQESTGWRPLNRYSHLET 187
Db 13 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 NTAPGEKYSLAYNOTAFILSCRTMYSLRRPKPHFESSLVAHFFE-----RE 280
QY 188 HAHLFKQAALPNGVKKASSPSPEPAVELSDSGQQEPED 226
Db 11 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 281 RVIILDACDAVISGVGVSSAK-----GTKHPRD 309

RESULT 16
P93012 ID P93012 PRELIMINARY; PRT; 182 AA..
AC AC P93012;
DT DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DI DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
GN GN Putative ubiquitin-conjugating enzyme E2.
OS OS ATG2G33770.
OC OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidis II; Brassicales; Brassicaceae; Arabidopsids.
ON NCBI_TaxID=3702;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLOMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounds S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Howman C.L., Harstead M.F., Feldblum T.V.,
RA Buell C.R., Kelchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanKen S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams W.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
```


RA M. Litscher, J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Kizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.B., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tait J., Lambuda G., Iorlami M.J., Town C.D.,
RA Usterback L., Van Aken S., Vaysberg M., Vyotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
RA "Sequence and analysis of chromosome 1 of the plant *Arabidopsis*
RA Thaliana".
RA Nature 408:816-826(2003).
RA EMBL: ACG19018; AAC52276.1;
RA HSP: P56550; 109A.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UBQ_conj_1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00238; BIR; 1.
DR PROSITE: PS50127; UBQ1; IN_CONJUGAT; 2; 1.
SQ SEQUENCE 244 AA; 27513 MW; D41C389F209C7C0B CRC64;

Query Match 28.5%; Score 374.5; DB 10; Length 244;
Best Local Similarity 40.8%; Pred. No. 4.3e-26;
Matches 76; Conservative 36; Mismatches 58; Indels 17; Gaps 4;

QY 21 MDIFSA...KGP:RPYPYDGLYLFQIQ:PNITYPAVPPHP:CYLSQSG:RLNPNLYDNGKVC 79
DB 1 MDL:RAVIAIGAESPYHUGLFFEDIOFDPDYSPVPPNYIHS---GGLRNPNLYDNGKVC 58
QY 80 VSLQIHWIKGKTERWTSK--SSLLQVLSIS:QGLT:WNEPYNFAGFDSDGLQFGYNSRC 138
DB 59 LSLIGLWHSNAQSM:PKESIMIQVLSV:QALV:NEQPYNFGYGLIKGTWIGKSKV 118
QY 135 YNEMALIRVQSMQTVRRPEVFEQIEQHFESTGGWRLVNRKESWLETHALLEKAQALP 198
DB 119 YSNVFLIS:KTMVYSMKPKQPHFRYVGNHY-----YVRSHDVIKACNAYK 165
QY 139 NGVPKAS 205
DB 166 AGAPIGS 172

RESULT 19
QYVH31
AC QYVH01 PRELIMINARY; PRT: 4904 AA.
DI 01-MAY-2000 (TRENUREL.13, Created)
DI 01-MAR-2001 (TRENUREL.16, Last sequence update)
DI 01-MAR-2002 (TRENUREL.20, Last annotation update)
DE C56303 protein.
GN BK0CF OR C66303.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Aphidrodia; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anandakis P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.E.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-B.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.J., Andrews Pfannkuch C., Haldwin D.,
RA Ballou R.M., Basu A., Haxendale J., Haytarkoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Blandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport I.B., Davies P.,
RA de Pablo H., Delcher A., Deng Z., Mays A.B., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fouts C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Godek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jalali M., Katush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.F., Kodira C.B., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Malte B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy R., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*."
RA Science 287:2185-2195(2000).
RL CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP -
CC DIPOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC EMBL: AE003686; AAF54520.2;
DR HSP: Q13490; IQBH.
DR FlyBase: FBgn0037808; Bruce.
DR InterPro: IPR001370; BIR.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00653; BIR; 1.
DR Pfam: PF00179; UBQ_conj; 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00238; BIR; 1.
DR PROSITE: PS50143; BIR_REPEAT_2; 1.
DR PROSITE: PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 4904 AA; 542137 MW; 626D616B6FE1A4CF CRC64;

Query Match 28.3%; Score 372.5; DB 5; Length 4904;
Best Local Similarity 33.0%; Pred. No. 3.5e-24;
Matches 98; Conservative 46; Mismatches 56; Indels 57; Gaps 10;

QY 4 LATS...EGIMVKTEDRMOLF:SGALIKGPRTPYEDGLYLFIDLPNTYPAVPP--H 57
DB 4572 LSTSLPLSFSSSVFVRCOTDRIDIMKVLITGPADITPYANGCFEVEFPDPY:NPQMLIN 4631
QY 58 FCYLSQSGRLNPNLYDNGKVCVSLGIGTWIGKTERWTSK--SSLLQVLSISQGLLVNRP 116
DB 4632 LETTGRIHSYRNPINLYDNGKVCVSLVNTWHRGPEEKWNAQTSSFLVLVLSIQSLIVPEP 4691
QY 117 YYNACGFDSDGLQFGYNSRCYNE---MALIRVQSMQTVRRPEVFEQIEQHFSTG 173
DB 4692 YFNPEGFERSRSPSGTSSREYNSNYQACVR--WAMLEQIRSPSOFKDVTKHFUK 4749
QY 174 GWRLVNRKESWLETHALLEKAQALP-----HALLEKAQALP--PNG----- 200
DB 4750 REEICQATGEMTEELKQYKPYTERKASMTISFNWMLVLRHRYHRELSKLPKPPGLGLIDA 4809
QY 201 -----VFKASSSEPPAVALESDSQCPEDGAP-----GEASQGSSEGAAG 246
DB 4810 PFNPVATLPMDVSAVAPSLA-ATNTVQAQPDADIVDTITLLGENADECEADGDAEG 4865

RESULT 20
Q8S193 PRELIMINARY; PRT: 493 AA.
ID Q8S193

AC Q9S193;
 DI 01-JUN-2002 (TrEMBLrel. 2.; Created)
 DI 01-JUN-2002 (TrEMBLrel. 2.; Last sequence update)
 DI 01-JUN-2002 (TrEMBLrel. 2.; Last annotation update)
 DE Putative ubiquitin-conjugating protein.
 GN B1144304.10.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Grizeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Susaki I., Matsumoto T., Yamamoto K.;
 RI "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
 RT clone:B1144304.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP003335; BAB90019.1; ...
 SQ SEQUENCE 493 AA; 55119 MW; R3C27B22264F1D0 CRC64;
 Query Match 28.3%; Score 372; DB 10; Length 493;
 Best local similarity 39.9%; Pred. No. 1.5e-25;
 Matches 81; Conservative 29; Mismatches 75; Indels 14; Gaps 2;
 QY 3 LIATLS:PEGIMVKTPEADMDLFSALIKGPTPTPYEDGLYLFDFDLQIPNTYPAVPHFC 62
 DB 253 VLEKLPDNIFRVVYEDKLELLRAVLIGPSTPYHSGLEFFDVYFPQYPRNPLIVYHS 312
 QY 63 QCSRLNPNIYDNGKVCVSLIGTWIGKIGRWTSKSSLLQVLIISOGILVNEPYNEAG 122
 DB 313 G-QRLNPNIYAGSAVCLSLINWFGCCCKWNPSTLQIVISQALVLAKEFYNEPG 371
 QY 123 FDSRGIQEGYENRCYNEMALIRVQSMITQLVRRPEVFEQEIQRHFTSGCWRLVNRIE 182
 DB 372 PESYANTPRAEKKSIAYNQETEL-SCKTMYLSLRNPKKHFDDFLIGHKHYG----- 423
 QY 183 SWLETHAL-EKQAQLPNGVPKAS 205
 DB 424 -----HSILIGCNAYMDQAQVGS 441
 RESULT 2;
 O98738
 ID O98738 PRELIMINARY; PRT: 4845 AA.
 AC O98738;
 DI 01-NOV-1998 (TrEMBLrel. 08; Created)
 DI 01-NOV-1998 (TrEMBLrel. 08; Last sequence update)
 DI 01-MAR-2002 (TrEMBLrel. 20; Last annotation update)
 DE Ubiquitin-conjugating enzyme.
 GN B1R36.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98292517; PubMed=9628897;
 RA Hauser H.P., Bardroff M., Pyrowolakis G., Jentsch S.;
 RT "A giant ubiquitin-conjugating enzyme related to IAP apoptosis
 R1 inhibitors.";
 RL J. Cell Biol. 141:1415-1422(1998).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: Y17267; CAA76720.1;
 DR HSSP: Q13493; 108H;
 DR MOP: MCI-1276108; B1R36.

DR InterPro: IPR001370; BIR.
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00653; BIR; 1.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: PD000461; UBQ_conjugat; 1.
 DR SMART: SM00238; BIR; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS0143; BIR_REPEAT_2; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ligase; Ubiquitin conjugation.
 SQ SEQUENCE 4845 AA; 528411 MW; 35C585F73B8CF84C CRC64;
 Query Match 28.3%; Score 372; DB 11; Length 4845;
 Best local similarity 40.8%; Pred. No. 3.9e-24;
 Matches 78; Conservative 35; Mismatches 68; Indels 10; Gaps 5;
 QY 4 LATSUP----EGIMVKTPEADMDLFSALIKGPTPTPYEDGLYLFDFDLQIPNTYPAVPHFC 59
 DB 4573 ISTSLPISSSSSVFCDEERIDIMKVLITGPDTPYANGCFEFDVYFPQYPSPP-IV 4631
 QY 60 YLSQCSG---RLAPNLYDNGKVCVSLIGTWIGKIGRWTSKSSLLQVLIISOGILVNE 115
 DB 4632 NLETTGGHSVRFNPVNDKVCLSILNTWIGRPEKWNPDTSFQLVLSVQSLILVAE 4691
 QY 116 PYNNAEDSDRGIQEGYENRCYN-EMALIRVQSMITQLVRRPEVFEQEIQRHFTSGG 174
 DB 4692 PYNEPQYRSRTPSGTOSREYDGNIRQATVKWALEQIRNPSPCFKEVHKHYEYK 4751
 QY 175 WELYNRIESML 185
 DB 4752 IELAAQCCHW 4762
 RESULT 22
 O988B7
 ID O988B7 PRELIMINARY; PRT: 1867 AA.
 AC O988B7;
 DI 01-MAR-2001 (TrEMBLrel. 16; Created)
 DI 01-MAR-2001 (TrEMBLrel. 16; Last sequence update)
 DI 01-MAR-2002 (TrEMBLrel. 20; Last annotation update)
 DE CDNA FLJ13786 fls, clone PLACE4000654, highly similar to Mus musculus
 DE ubiquitin conjugating enzyme.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Tsogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
 RT "NEVO human cDNA sequencing project.";
 KL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AK023848; BAB14700.1;
 DR HSSP: P15731; 10CQ;
 DR InterPro: IPR000345; CytC_hcmc_bind.
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: PD000461; UBQ_conjugat; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ligase; Ubiquitin conjugation.

```

DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBCC; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Hypothetical protein; Ligase; Ubiquitin conjugation.
SQ SEQUENCE 246 AA; 28075 MW; A0A668FB9430ACDB CRC64;

Query Match      18.9%; Score 248; DB 4; Length 246;
Best Local Similarity 43.7%; Pred. No. 1.4e-14;
Matches 59; Conservative 16; Mismatches 50; Indels 10; Gaps 5;

QY   9 PEGIMVKTFEDRMDLFSALIKGPTTRIPYEDGLYLFDIQLPNTIYPAPVPHFCYLSQCSC-- 66
Db    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Dy   8 PPGMEVPDVDMTKIHAIITGPDTPYEGGFELFVERCPDPPIHPERVKLTMTGNNTV 67
QY   57 RLNPVLNDNGKVCVSLIGTWICKGTERTWTSKSSLIQVLSIOGLILVNPEYYNAGCFUSD 126
Db    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Dy   68 RPNPNFYNGKVCLSIIGTWTGPA---WSPAQSISSVLSIOSLMTEN-PYNHEPGFEQE 123
QY   127 RGLQEGYENSRCYNE 141
Db    | - | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Dy   124 R--HPG--DSKNYNE 134

RESULT 24
Q8RZT4 PRELIMINARY; PRT: 246 AA.
ID Q8RZT4;
AC Q8RZT4;
DI 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DI 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 28.1 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBTaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.:
RL Submitted (APR-2002) to the EMBL/GenBank/DDAJ databases.
DR FMRL: BC027266; AAAH27266.1;
SQ Hypothetical protein.
KW SEQUENCE 246 AA; 28075 MW; A0A668FB9430ACDB CRC64;

Query Match      18.9%; Score 248; DB 11; Length 246;
Best Local Similarity 43.7%; Pred. No. 1.4e-14;
Matches 59; Conservative 16; Mismatches 50; Indels 10; Gaps 5;

QY   9 PEGIMVKTFEDRMDLFSALIKGPTTRIPYEDGLYLFDIQLPNTIYPAPVPHFCYLSQCSC-- 66
Db    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Dy   8 PPGMEVPDVDMTKIHAIITGPDTPYEGGFELFVERCPDPPIHPERVKLTMTGNNTV 67
QY   67 RLNPVLNDNGKVCVSLIGTWICKGTERTWTSKSSLIQVLSIOGLILVNPEYYNAGCFUSD 126
Db    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Dy   68 RPNPNFYNGKVCLSIGTWTGPA---WSPAQSISSVLSIOSLMTEN-PYNHEPGFEQE 123
QY   127 RGLQEGYENSRCYNE 141
Db    | - | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Dy   124 R--HPG--DSKNYNE 134

RESULT 25
Q9H893 PRELIMINARY; PRT: 236 AA.
ID Q9H893;
AC Q9H893;
DI 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DI 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE THYKO1000983 protein (Hypothetical 27.0 kDa protein).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBTaxID=9606;
RN [1]
```

RA	Knehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sasaki K., Okido T., Furuno M., Aono H., Baldarelli K., Harsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Buit C., Fletcher C., Fujita M., Cariboldi M., Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Nombarts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Soya T.I., Shibata Y., Storch K.F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Horris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S., Hayashizaki Y.;
RT	*Functional annotation of a full-length mouse cDNA collection.*;
RL	Nature 409:685-690(2001).
CC	-!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC	PROTEINS (BY SIMILARITY).
CC	-!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC	DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC	-!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC	-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC	THIOLESTER FORMATION (BY SIMILARITY).
CC	-!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR	EMBL; AK005524; BAB24097.1; -;
DR	HSSP; Q02159; ZUCZ.
DR	InterPro; IPR000608; UBQ_conjugat.
DR	Pfam; PF06179; UQ_con; 1_conjugat.
DR	ProDom; PD000461; UBQ_conjugat; 1.
DR	SMART; SM00212; UHC; 1.
DR	PROSITE; PS01017; UBIQUITIN_CONJUGAT_2; 1.
KW	Ligase; ubiquitin conjugation.
QY	SEQUENCE 232 AA; 26459 MW; 2279DAFD393A59A CRC64;
Query Match 18.4%; Score 242; DB 11; Length 232;	
Best Local Similarity 46.6%; Pred. No. 4.6e-14;	
Matches 55; Conservative 15; Mismatches 38; Indels 10; Gaps	
Qy	26 ALIKGRTPTPEYGLYLFDIQLGNIVPAVPHFCYLSQSG--RLNPNLYDNCKVCVSL 83
Db	: : : : : : :
Db	11 ALITGPDTPYGGFFLVKFCPPYPIHPKVLMTIGNTVFNPFPNCKVCLSL 70
Qy	84 GTWGTGTERWTSKSLLOWLISQGLILVNPYPNPFAGSDRGLQRYENSRVCNE 141
Db	: : : : : : : : : :
Db	71 GTWTGPA--WSPASISVSLISIQSLMTEN-PVHPNPGFEQER--HPG--DSKNYNE 120
RESULT 27	
Q9BTCL	PRELIMINARY: PRT: 222 AA.
ID	Q9BTCL;
AC	01-JUN-2001 (TREMBLrel. 17, Created)
DI	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE	Ubiquitin carrier protein.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX	NCBI_taxid=9606;
KN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=COLON;
RA	Strausberg R.;
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC	-!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC	PROTEINS (BY SIMILARITY).
CC	-!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC	DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC	-!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC	-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-

```

CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: BC004236; RAH04236.1; -.
DR EMBL: BC007554; AAK07554.1; -.
DR HSSP: P15731; 100%.
DR InterPro: IPR000608; UQ conjugat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; UQ_conjugat; 1.
DR SMART: SM00212; UQCC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 222 AA; 23845 MW; 2842DC3DCD2AFCB5 CRC64;

Query Match 12.9%; Score 169; DB 4; Length 222;
Best Local Similarity 24.5%; Pred. No. 1.9e-07;
Matches 60; Conservative 36; Mismatches 73; Indels 76; Gaps 10;

CY 4 LATSIPGIMVKTFFEDRMDFLSALIKGPTRTPYEDGILYLFQIQ.PNIYPVAVPHFCYLSQ 63
DB 23 LTASPPGKIKVFNEEDLIDQVTEGPEGIPYAGLFEMKLLKGLKDFPASPCKGYELTK 82
CY 64 CSGRAPNLYDNGKVCVSLLSGKGTERTWTSKSSLLQVLSISQGLILVNEPYNEAGF 123
DB 83 L---FHPNVCANGRCVNVL-----KRWTAELGIRKVLTKCLLHPNP---FSAL 129
CY 124 QSDRG--LQGTENSRCYNEMALIRVQSMTOVLVRPPVEFQEIRQHFSTGWRILVNR 181
DB 130 NEAGRLLEENEE---YAARARL-----LTEL-----HGGAGG----- 160
CY 182 ESWLETHALLEKAQALPNCVFKASSPEPPAVAKLSDCOEPDGGPARGASQSDSE 241
DB 161 -----PSGRARASRALSGIEASLIDGAP-----GGP----- 188
CY 242 GAGG 245
DB 185 GAGG 193

RESULT 28
Q9XHP3 PRELIMINARY: PRT; 153 AA.
AC Q9XHP3;
DI 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, last sequence update)
DE 01-MAR-2002 (TRENBLrel. 25, last annotation update)
DE Ubiquitin-conjugating enzyme E2.
GN UBIE2.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
OC Vinceae; Catharanthus.
OX NCBI_TaxID=4058;
RN 1;
RP SEQUENCE FROM N.A.
RA Gante P., Nepomuceno G., Siberil Y., Thiersault M., Boisson B.,
RA Doireau P.;
RI "Molecular cloning of a Catharanthus roseus cDNA encoding ubiquitin
RI conjugating enzyme E2."
RI Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: AF091621; AAG42941.1; -.
DR HSSP: P15731; 100%.
DR InterPro: IPR000608; UQ conjugat.
DR Pfam: PF00179; UQ_con; 1.

```

```

DR ProDom: PD000461; UQ_conjugat; 1.
DR SMART: SM00212; UQCC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 153 AA; 17140 MW; 367FF982H5B20H83 CRC64;

Query Match 12.8%; Score 168; DB 10; Length 153;
Best Local Similarity 28.3%; Pred. No. 1.4e-07;
Matches 43; Conservative 27; Mismatches 62; Indels 20; Gaps 4;

QY 3 LIATSLPGIMVKTFFEDRMDFLSALIKGPTRTPYEDGILYLFQIQ.PNIYPVAVPHFCYLS 62
DB 17 LISEFAP-GISASPSDDNMRFNVMLGPAQSPYEGCVKLELFLEPRYPNAPKVRFLT 75
QY 63 CCSGRAPNLYDNGKVCVSLLSGKGTERTWTSKSSLLQVLSISQGLILVNEPYNEAG 122
DB 76 KI---YHPNIDKILGRICLDIL-----KKWSPALQIRTVLISIQALLSAPND----- 119
QY 123 FDSDRGLQEGVNSRCYNEMALIRVQSMTOVL 154
DB 120 ---DDPLESENIAKUNKSNEAEAVETAKENTRL 148

RESULT 29
Q94A97 PRELIMINARY: PRT; 153 AA.
AC Q94A97;
DI 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, last sequence update)
DE 01-MAR-2002 (TRENBLrel. 20, last annotation update)
DE Atlg78870/F9K20.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 1;
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida G., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin C., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RI "Arabidopsis cDNA clones."
RI Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: AY049261; AAK83603.1; -.
DR InterPro: IPR000608; UQ_conjugat.
DR Pfam: PD000461; UQ_conjugat; 1.
DR ProDom: PD000461; UQ_conjugat; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; UNKNOWN_1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 153 AA; 17192 MW; 66FDD14FA347A800 CRC54;

Query Match 12.6%; Score 166; DB 10; Length 153;
Best Local Similarity 28.3%; Pred. No. 2.2e-07;
Matches 43; Conservative 27; Mismatches 62; Indels 20; Gaps 4;

QY 3 LIATSLPGIMVKTFFEDRMDFLSALIKGPTRTPYEDGILYLFQIQ.PNIYPVAVPHFCYLS 62
DB 17 LLSFPAP-GISASPSDDNMRFNVMLGPAQSPYEGCVKLELFLEPRYPNAPKVRFLT 75

```

QY 63 QCSGRNPNLYDNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLLSIQGLILVNEPYPYNEAG 122
 Db 76 KT---YIPNTDKLGRICLDL-----KDKSPALQIRIVLLSIQALLSAPNP----- 119
 QY 123 FDSNGLQCEGYENRCYNEMALIRVQSMTQL 154
 Db 120 ---DDPLSENIAKHWSNEAFVETAKETRL 148

Query Match 12.5% Score 164; DR 10; Length 153;
 Best Local Similarity 27.6%; Pred. No. 3.3e-07;
 Matches 42; Conservative 28; Mismatches 62; Indels 20; Gaps 4;

QY 3 LLATSLPEGIWVKTFFEDRMDLFSALIKGPIRTPYENGILYFDLQLPNIYPAVPEHFCYLS 62
 Db 17 LLSEAP-CISASPSFENMRYFNMILCPTQSPGYGFKLEFLREYFMAAPKVRFL 75
 QY 63 QCSGRNPNLYDNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLLSIQGLILVNEPYPYNEAG 122
 Db 76 KT---YIPNTDKLGRICLDL-----KDKSPALQIRIVLLSIQALLSAPNP----- 119
 QY 123 FDSNGLQCEGYENRCYNEMALIRVQSMTQL 154
 Db 120 ---DDPLSENIAKHWSNEAFVETAKETRL 148

RESULT 31
 Q9NPD8 PRELIMINARY; PRI: 197 AA.

AC Q9NPD8
 DT 01-OCT-2003 (TrEMBLrel. 15, Created)
 DT 01-OCT-2003 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Ubiquitin-conjugating enzyme ISOLOC (HSPC150) (Ubiquitin-conjugating
 DE enzyme E2) (CDNA FLJ20497 FIS, clone KAT08890) (HSPC150 protein
 DE similar to ubiquitin-conjugating enzyme).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okazaki H., Hayashi A., Kozuma S., Salto T.;
 RA "Ubiquitin-conjugating enzyme isolog.";
 RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ye M., Zhang Q.H., Zhou Y., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
 RT "Human full length cDNA cloned from cd34+ stem cells.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ADRENAL GLAND;
 RA Huang C., Zhang C., Wu T., Peng Y., Gu Y., Zhang L., Jiang C., Li Y.,
 RA Han Z., Wang Y., Chen Z., Fu G.;
 RT "A novel gene expressed in human adrenal gland.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Watanabe K., Kumaoki A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Oabayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isogai T., Sugano S.;
 RT "NMDA human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LUNG;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LUNG;
 RA Strausberg R.;
 RL Submitted (JFC-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER

QY 63 QCSGRNPNLYDNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLLSIQGLILVNEPYPYNEAG 122
 Db 76 KT---YIPNTDKLGRICLDL-----KDKSPALQIRIVLLSIQALLSAPNP----- 119
 QY 123 FDSNGLQCEGYENRCYNEMALIRVQSMTQL 154
 Db 120 ---DDPLSENIAKHWSNEAFVETAKETRL 148

Query Match 12.5% Score 164; DR 10; Length 153;
 Best Local Similarity 27.6%; Pred. No. 3.3e-07;
 Matches 42; Conservative 28; Mismatches 62; Indels 20; Gaps 4;

QY 3 LLATSLPEGIWVKTFFEDRMDLFSALIKGPIRTPYENGILYFDLQLPNIYPAVPEHFCYLS 62
 Db 17 LLSEAP-CISASPSFENMRYFNMILCPTQSPGYGFKLEFLREYFMAAPKVRFL 75
 QY 63 QCSGRNPNLYDNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLLSIQGLILVNEPYPYNEAG 122
 Db 76 KT---YIPNTDKLGRICLDL-----KDKSPALQIRIVLLSIQALLSAPNP----- 119
 QY 123 FDSNGLQCEGYENRCYNEMALIRVQSMTQL 154
 Db 120 ---DDPLSENIAKHWSNEAFVETAKETRL 148

RESULT 30
 Q9FZ48 PRELIMINARY; PRI: 153 AA.

AC Q9FZ48
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative ubiquitin-conjugating enzyme E2 (At.g16890/F17P16.16).
 GN F17P16.16
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Federspiel N.A., Palm C.J., Conway A.B., Gunn L., Hansen N.F.,
 RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
 RA Buchler E., Chao Q., Chao C., Chao J., Chai E., Gonzalez A.,
 RA Lewis B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
 RA Leuz C., Liu A., Liu S., Mikharzky N., Pham P., Sakano H., Shinn P.,
 RA Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cheek R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Roeser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Shida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Ouedera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
 RA Roeser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Shida J., Jones J., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Ouedera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN -> PROTEIN-LYSINE-AMP +
 CC DIPHOSPHATE + PROTEIN-N-UBIQUITYLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AC051629; AAF9844.1; -;
 DR EMBL: AY052407; AAK9500.1; -;
 DR EMBL: AY061926; AAL31253.1; -;
 DR HSSP: P15731; 1000.
 DR InterPro: IPR000608; GQO_conjugat.
 DR Pfam: PF001179; UQ_con: 1.
 DR ProDom: PD000461; UQO_conjugat: 1.
 DR SMART: SM0212; UHCC: 1.

CC PROTEINS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +

CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.

CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.

CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-

CC THIOL-ESTER FORMATION (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.

DR EMBL: A0032931; BA093711.1; -

DR EMBL: A161449; AAF29114.1; -

DR EMBL: AF160215; AAF67016.1; -

DR EMBL: AK000504; AAF91211.1; -

DR EMBL: BC094152; AAU04152.1; -

DR EMBL: BC019284; AAU05284.1; -

DR HSP: P15731; 1003.

DR JctPr: IPR000608; UQO_conjugat.

DR Pfam: PF00179; UQ_con; 1.

DR ProDom: PD000461; UQO_conjugat; 1.

DR SMART: SM0012; UQCC; 1.

DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.

DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.

KW Ligase; Ubiquitin conjugation.

SQ SEQUENCE 197 AA; 22521 MW; 60320774A7FA528A CRC64;

Query Match 12.5%; Score 164; DB 4; Length 197;

Best Local Similarity 28.3%; Pred. No. 4.6e-07;

Matches 45; Conservative 33; Mismatches 75; Indels 16; Gaps 6;

QY 3 LKATSLPGKIVKIFEDHMDLFSALIKGTRIPYEDGLYFDIQLNLYPAVPPHPCYLS 62

DB 13 MLEATPEPGITCWSKCOMDRAQILUGANIPYEGVKIEVILPERYEPFQIRELT 72

QY 63 QCSGRNPALDNGKVCVSLIGTWIGKTERWTSKSSLLQVLISIQGLIIVNEP 122

DB 73 P1---VNPIDNSGRICLDVL-KLPPKA--WPPSLNATVLISIQ--LMSKP----- 118

QY 123 FDSRGLQKYSKCYNEMALIKVQSWTQL-VRRPEVFEQERQHFSTGG 174

DB 119 -NPEEDPLADISSEFKYKPAFLANQWTERHAKOKAKADEEMLDLPKAG 170

RESULT 32

QSR17 ID QSR17 PRELIMINARY; PRI: 162 AA.

AC QSR17;

DI 01-JUN-2002 (Tremblrel. 21, Created)

DI 01-JUN-2002 (Tremblrel. 21, Last sequence update)

DI 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Ubiquitin conjugating enzyme E2-17kDa.

GN E2010.13.01.

OS Eucephalitozoan cuniculi.

CC Eukaryota; Microsporidia; Unikaryonidae; Eucephalitozoan.

OX NCBI_TaxID:6035;

RN 1;

RP SEQUENCE FROM N.A.

RC STRAIN-GB-M1;

RA Genoscope;

RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

RN 12;

RP SEQUENCE FROM N.A.

RC STRAIN-GB-M1;

RX MEDLINE-21576510; PubMed-11719806;

RA Katinka M.D., Duprat S., Corfiliot E., Melenier G., Thomarar F.,

RA Vronster G., Harbe V., Peyretaillet E., Brottier P., Winkler P.,

RA Leibaic F., El Alaoui H., Peyret P., Saurin W., Gouy M.,

RA Weissenbach J., Vivares C.P.;

RI *Genome sequence and gene compaction of the eukaryote parasite

RI Eucephalitozoan cuniculi.*

RL Nature 414:450-453(2001).

DR EMBL: AL590449; CA025850.1; -

SQ SEQUENCE 162 AA; 18305 MW; C03FF20663E2408B CRC64;

Query Match 12.4%; Score 163; DB 5; Length 162;

Best Local Similarity 33.3%; Pred. No. 4.4e-07;

Matches 35; Conservative 22; Mismatches 36; Indels 12; Gaps 3;

QY 12 LKATSLPGKIVKIFEDHMDLFSALIKGTRIPYEDGLYFDIQLNLYPAVPPHPCYLS 71

DB 24 IFAPLEDDLMIAAVIFGPDSTPEEGTFLSVLTFFHYDQPPAVREVSE---MFRPN 80

QY 72 LDNGKVCVSLIGTWIGKTERWTSKSSLLQVLISIQGLIIVNEP 116

DB 81 IYPNGHCLDIL-----SNRNPSYDVIGVLISIQSLI--NDP 116

RESULT 33

Q9C2A5 ID Q9C2A5 PRELIMINARY; PRI: 166 AA.

AC Q9C2A5;

DI 01-JUN-2001 (Tremblrel. 17, Created)

DI 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DI 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Probable ubiquitin-conjugating enzyme ubcP3.

GN B112.230.

OS Neurospora crassa.

CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

CC Sordariales; Sordariaceae; Neurospora.

OX NCBI_TaxID:5141;

RN 1;

RP SEQUENCE FROM N.A.

RA Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann R., Holland R.,

RA Nyakatura G., Mewes H.W., Mannhaupt G.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.

RN 12;

RP SEQUENCE FROM N.A.

RA German Neurospora genome project;

RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.

CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER

CC PROTEINS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +

CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.

CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.

CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-

CC THIOL-ESTER FORMATION (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.

DR EMBL: AL513444; CAC28704.1; -

DR HSP: Q02159; 2UCZ.

DR InterPro: IPR000608; UQO_conjugat.

DR Pfam: PF00179; UQ_con; 1.

DR ProDom: PD000461; UQO_conjugat; 1.

DR SMART: SM0012; UQCC; 1.

DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.

DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.

DR Ligase; Ubiquitin conjugation.

KW SEQUENCE 166 AA; 18540 MW; 2387742CA5477185 CRC64;

SQ SEQUENCE 166 AA; 18540 MW; 2387742CA5477185 CRC64;

Query Match 12.3%; Score 162; DB 3; Length 166;

Best Local Similarity 31.3%; Pred. No. 5.6e-07;

Matches 36; Conservative 26; Mismatches 43; Indels 10; Gaps 3;

QY 4 LKATSLPGKIVKIFEDHMDLFSALIKGTRIPYEDGLYFDIQLNLYPAVPPHPCYLS 62

DB 17 LTNNPPEGTAGPTISDDLLHWECLIQCPGPIFGGVEPAEEKFPNDYHPHPTMKFLG 76

QY 63 QCSGRNPALDNGKVCVSLI-----GTWICKGTERTWTSKSSLLQVLISIQGLI 111

DB 77 DI---FHPNVPYSGVLVCISILHPPGDDPNHYETASEHNSPISQSVKILISVMSML 128

RESULT 34

Q96FX4 ID Q96FX4 PRELIMINARY; PRI: 152 AA.

AC Q96FX4;

DI 01-DEC-2001 (Tremblrel. 19, Created)

DI 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DI 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Ubiquitin-conjugating enzyme E2A (RAD6 homolog).

DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; i.
DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; i.
KW Tagase; Ubiquitin conjugation;
SQ SHOUENGE 152 AA; 17315 MW; UAAB5B7770E47E2 CRC64;

Query Match 12.2%; Score 160.; DB 13; Length 152;
Best Local Similarity 29.6%; Pred. No. 7.5e-07;
Matches 42; Conservative 27; Mismatches 59; Indels 14; Gaps 4;

QY 9 PEGIMVKTFEDRMOLFSAIKGPTRIPEYDGLFDIQLPNIYPVAPPHCYLSOCSGML 68
| | : : : : | | | | | | : : : : | | : : |
Db 21 PAGVSASSENIMWNNVFGPEDEGTFLKIETEEYPNKPTVEFSK---HF 77
| | : : : : | | : : : : | | : : : : | | : : |
QY 69 NFNLYDNKKVCVSLGWIGTKTERTWKSSLLQSIVLISITGLIDNVEPYNEAGFSDRG 128
| | | : | : | : | : | : : : : | | : : : : | | : : |
Db 78 HFNVTADGSICDIL-----QNWSPTVDVSSIIITQSLL--DEGNPNSPA--NSQA 126
| | : : : : | | : : : : | | : : : : | | : : |
QY 125 LGGGYENSKYNEMALIRVVQS 150
| | : : : : | | : : : : | | : : : : | | : : |
Db 127 AOLYOENREYKRVSALVEQS 148
| | : : : : | | : : : : | | : : : : | | : : |

RESULT 37
QCUC37 PRELIMITARY; PRT; 204 AA.

ID QCUC37
AC QCUC37
DT 01-JUN-2001 (1RBMJurel). 17, Created)
D: 01-JUN-2001 (1RBMJurel). 17, Last sequence update)
U: 01-MAR-2002 (1RBMJurel). 20, Last annotation update)

DF 2700084122Rik protein.
GN 2700084122Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mamalia; Eutheria; Rodentia; Soturognathi; Muridae; Mus.
ON NCBI_Taxid:10090;
RX [1]
RP SEQUENCE FROM N.A.

KC MEDLINE-CSTRL/53; TISSUE-EMBRYONIC STEM CELLS, AND EMBRYO;
KM STRUTINE-21085660; PubMed11217851;
KA Kawai J., Shingawa A., Shibata K., Yoshino H., Adachi J., Fukuda S.,
RA Arakawa I., Hara A., Fukunishi Y., Kosho M., Itoh M., Ishii Y.,
YA Mizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,
SA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
FA Fritschmann W., Gaasterland T., Gissi C., King H., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anjo H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Donato M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
GA Gastineich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni I., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring H., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Teyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
HA Hayashizaki Y.;
FT Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).

CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
DIAPHOSPHATE + PROTEIN N-UBIQUITYLVISINE.
CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
THIOESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL; AK021213; BAB32332.1; .
DR FMHI; AK012565; BAH28120.1; .
DR HSP; p15731; LOC.
DR MGD; MG:1914446; 2700084122Rik.
DR InterPro; IPRO00608; UBQ_conjugat.
DR Pfam; PF00179; UQ_con; i.

Query Match 12.1% Score 158.5; DB 3; Length 151;
 Best Local Similarity 28.3%; Pred. No. 1.4e-06;
 Matches 39; Conservative 24; Mismatches 54; Indels 21; Gaps 4;

QY 4 LATSIPGIMVKTFTEDMDLSALIKGTRTPYEDGLYFDIQLPNIPYAPVPHFYLSQ 63
 DB 16 MQIDPPAGYASAPVENVMTWNAVIGPAUTPFEDGTFRVWOFEOYXPKPKPVKFI 75
 QY 64 CSGRLNPNLYDKKVCVSLGATGWICKGRTWTSKSSLLQVLISIOGLILVNEPYNEAGF 123
 DB 76 ---MFHFNVTATGELCLOIL-----QNRKSTPYDVAVALTISQSLI--NDP----- 116
 QY 124 QSDHGLQECYENSCYNE 141
 DB 117 ---NRGSTANVEASNLTK 132

RESULT 39
 Q9VX25 PRELIMINARY: PRI: 190 AA.
 AC Q9VX25;
 DT C2-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE C88188 protein.
 CN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Polygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RF SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE 20196006; PubMed-10731132;
 RA Adams M.D., Celinker S.E., Holt K.A., Evans C.A., Gocayne J.D.,
 RA Ananathan P.G., Scher S.E., Li P.W., Hoskins R.A., Gallo R.F.,
 RA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan H.H., Doyle C., Baxter E.G., Holt G., Nelson G.R., Miklos G.L.G.,
 RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bencs P.V., Bertram R.P., Bhandari D., Boishakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paulis R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Duan-Rocha S., Dunkov B.C., Durn P.,
 RA Durbin K., Ewing R., Garg N.S., Gerhart J., Giannone S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland J.J., Wei M.-H., Iqbalwala C.,
 RA Jalali M., Kalish J., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasbo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard C., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Sanders K.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidon-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Stadling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter F., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of Drosophila melanogaster."
 Science 287:2185-2195(2000).

CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN-N-UBIQUITYLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AE003506; AAF48756.1; -;
 DR HSSP: P15731; 1000.
 DR FlyBase: FBgn0030863; CG8188.
 DR InterPro: IPR002453; Beta_tubulin.
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00179; UBQ_con: 1.
 DR ProDom: PD000461; UBQ_conjugat; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ligase; ubiquitin conjugation.
 SQ SEQUENCE 190 AA; 21174 MW; 80F50C2F11272526 CRC64;

Query Match 12.1% Score 158.5; DB 5; Length 150;
 Best Local Similarity 28.0%; Pred. No. 1.4e-06;
 Matches 44; Conservative 34; Mismatches 56; Indels 23; Gaps 6;

QY 4 LATSIPGIMVKTFTEDMDLSALIKGTRTPYEDGLYFDIQLPNIPYAPVPHFYLSQ 63
 DB 7 METTPGEGIKVLINESDVTIOALDGPAGTPYAGATFRVKLTNKLKPLTPPKAYELTK 66
 QY 64 CSGRLNPNLYDKKVCVSLGATGWICKGRTWTSKSSLLQVLISIOGLILVNEPYNEAGF 123
 DB 67 I---FHNVAANGECVNTL-----KKWKPDIGTKHILUTIKCLLIVPNP---PSAL 113

QY 124 DSDRG--IOEYSYENSCYNEALIRVQSMITQLVRRP 158
 DB 114 NEEAGKMLIERYDQ---YSGRA----RMMEIHAQP 142

RESULT 40
 Q95XN7 PRELIMINARY: PRI: 329 AA.
 AC Q95XN7;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical 37.4 kDa protein.
 GN Y71G12B.15.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 CC Rhabditidae; Pelodcrinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-99069613; PubMed-9851916;
 RA None;
 RA "Genome sequence of the nematode C. elegans: a platform for
 RA investigating biology. The C. elegans Sequencing Consortium."
 RA Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Du H., Lamar B., Wohldmann P.;
 RA "The sequence of C. elegans cosmid Y71G12B."
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RA "Direct Submission."
 RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).

```

CC -|- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE
CC -|- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -|- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: AC025726; AAK73914.2; -
DR InterPro: IPR006098; UBO_conjugat.
DR Pfam: PF00179; UQ_con; -
DR PRODOM: PD000461; UBO_conjugat; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; UNKNOWN_1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Hypothetical protein; Ligase; Ubiquitin conjugation.
SQ SEQUENCE 329 AA: 37378 MW: 553094F6/3B2F109 CRC64:

Query Match 12.08; Score 157.5; DB 5; Length 329;
Best Local Similarity 31.38; Pred. NO. 3.5e-06;
Matches 35; Conservative 26; Mismatches 41; Indels 11; Gaps 3;

QY 10 EGIMVKTFELRMFLSALIKGPIETPYEDGLYLFQIPNIYPVAPPHFCYLSQCSGRLN 69
DB 114 KGFICVWELNLFVWLVGIVGPKLYQGGYFKASIFPSNYPSPSMKTTIKV---MH 170
QY 70 PNYKDKKVCVSLGTTWIK-----GIERWTSKSSLQVLISIQGLVNEP 116
DB 171 PNWYENGDLGISLHSPIDPQSGELACERWNPQSVRTILSV--ISLLNEP 221

Search completed: Apr. 10, 2003, 10:36:10
Job time : 51.7056 secs

```

Genere version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:32:32 : Search time 10.294: Seconds
(without alignments)
1484.732 Million cell updates/sec

Title: us-09-930-026-1

Perfect score: 1315

Sequence: 1 MALLATSIPEGIMVKTFDR.....PGEASGSDSFQGAQGIATFS 250

Scoring table: BLUSCM62

Gapop 10.0, Gapext 0.5

Searches: 248812 seqs, 6136940 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum db seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 135 summaries

Database : Published Applications, AA:*

- 1: /cgn2.6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2.6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2.6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2.6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2.6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2.6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2.6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2.6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2.6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2.6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2.6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2.6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2.6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2.6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	370	28.1	805	10	US-09-799-777-19
2	248	18.9	252	9	US-09-989-920-204
3	169	12.9	283	10	US-09-925-301-979
4	165.5	12.6	165	9	US-10-157-869-33
5	164	12.5	197	9	US-09-932-941-1904
6	164	12.5	197	9	US-09-849-626-1904
7	164	12.5	197	9	US-10-117-754-1904
8	160	12.2	204	9	US-10-180-927-2
9	157	11.9	196	10	US-09-925-301-868
10	156	11.9	232	9	US-10-122-806-446
11	148	11.3	288	9	US-10-060-019-23
12	145.5	11.1	165	10	US-09-842-528-4
13	140.5	10.7	144	10	US-09-842-528-5
14	137.5	10.5	165	9	US-10-157-869-17
15	134	10.2	147	9	US-10-108-605-331
16	130	9.9	143	9	US-10-043-487-392
17	130	9.9	143	9	US-10-043-487-399
18	130	9.9	148	10	US-09-926-312-4
19	127	9.7	157	10	US-09-785-671-5

20	118	9.0	158	10	US-09-785-671-2	Sequence 2, Appl
21	117	9.0	205	10	US-09-925-302-525	Sequence 525, App
22	117	8.9	170	10	US-09-976-165-22	Sequence 22, Appl
23	109	8.3	295	9	US-10-060-019-22	Sequence 22, Appl
24	107.5	8.2	153	9	US-10-043-487-394	Sequence 394, App
25	96	7.3	153	9	US-10-043-487-395	Sequence 395, App
26	95.5	7.3	176	9	US-10-043-487-401	Sequence 401, App
27	94.5	7.2	186	10	US-09-925-302-746	Sequence 746, App
28	94.5	7.2	305	10	US-09-925-302-620	Sequence 620, App
29	94	7.1	185	9	US-10-157-669-1	Sequence 1, Appl
30	94	7.1	185	10	US-09-840-787-2	Sequence 2, Appl
31	92.5	7.0	156	9	US-10-043-487-412	Sequence 412, App
32	92.5	7.0	256	10	US-09-925-301-1083	Sequence 1083, App
33	88.5	6.7	384	9	US-09-765-0618-72	Sequence 72, Appl
34	85.5	6.5	190	9	US-10-102-806-723	Sequence 723, Appl
35	84	6.4	392	9	US-09-765-0618-77	Sequence 77, Appl
36	83	6.3	151	10	US-09-842-528-2	Sequence 2, Appl
37	83	6.3	161	10	US-09-925-297-801	Sequence 801, App
38	82.5	6.3	434	10	US-09-879-957-22	Sequence 22, Appl
39	81.5	6.2	384	9	US-09-765-0618-73	Sequence 73, Appl
40	80	6.1	180	9	US-10-157-669-32	Sequence 32, Appl
41	79	6.0	2890	10	US-09-815-242-11410	Sequence 11410, A
42	79	6.0	2890	10	US-09-815-242-11571	Sequence 11571, A
43	78.5	6.0	328	10	US-09-900-527-2	Sequence 2, Appl
44	78.5	6.0	372	9	US-09-765-0618-78	Sequence 78, Appl
45	77.5	5.9	414	10	US-09-802-213-3	Sequence 3, Appl
46	77.5	5.9	414	10	US-09-925-297-747	Sequence 747, App
47	77.5	5.9	546	10	US-09-879-957-14	Sequence 14, Appl
48	77.5	5.9	655	9	US-10-054-683-29	Sequence 29, Appl
49	77.5	5.9	655	12	US-10-001-632A-2	Sequence 2, Appl
50	77.5	5.9	746	9	US-10-142-356-8	Sequence 8, Appl
51	77.5	5.9	873	9	US-10-200-154-2	Sequence 2, Appl
52	77.5	5.9	873	10	US-09-954-043-2	Sequence 2, Appl
53	77	5.9	404	10	US-09-879-957-24	Sequence 24, Appl
54	76	5.8	527	9	US-09-854-133-216	Sequence 216, App
55	76	5.8	527	10	US-09-738-973-216	Sequence 216, App
56	76	5.8	846	10	US-09-815-242-13904	Sequence 13904, A
57	75.5	5.7	372	9	US-09-765-0618-74	Sequence 74, Appl
58	75.5	5.7	479	10	US-09-925-302-527	Sequence 527, App
59	75.5	5.7	1332	10	US-09-982-091A-4	Sequence 4, Appl
60	75	5.7	496	10	US-09-815-242-10714	Sequence 10714, A
61	75	5.7	2000	12	US-10-010-901-29	Sequence 29, Appl
62	74.5	5.7	344	10	US-09-925-301-967	Sequence 967, App
63	74.5	5.7	371	9	US-10-025-380-1109	Sequence 1109, App
64	74.5	5.7	371	10	US-09-922-217-1209	Sequence 1209, App
65	74.5	5.7	406	10	US-09-925-301-1049	Sequence 1049, App
66	74.5	5.7	810	9	US-09-712-363-281	Sequence 281, App
67	74	5.6	334	10	US-09-755-830-43	Sequence 43, Appl
68	74	5.6	416	9	US-10-067-761-13	Sequence 13, Appl
69	74	5.6	416	10	US-09-804-156-13	Sequence 13, Appl
70	74	5.6	461	10	US-09-755-830-38	Sequence 38, Appl
71	74	5.6	516	10	US-09-740-046-4	Sequence 4, Appl
72	74	5.6	1356	9	US-09-969-037-7	Sequence 7, Appl
73	74	5.6	1356	9	US-10-022-939-2	Sequence 2, Appl
74	74	5.6	1356	9	US-10-100-405A-2	Sequence 2, Appl
75	73.5	5.6	472	10	US-09-745-763-36	Sequence 36, Appl
76	73.5	5.6	473	9	US-09-984-271-120	Sequence 120, App
77	73.5	5.6	640	9	US-09-983-204-16	Sequence 16, Appl
78	73.5	5.6	640	9	US-10-133-157-5	Sequence 5, Appl
79	73	5.6	305	10	US-09-734-569-158	Sequence 158, App
80	73	5.6	347	10	US-09-816-127-3	Sequence 3, Appl
81	73	5.6	394	10	US-09-925-301-1388	Sequence 1388, App
82	73	5.6	416	10	US-09-815-876-2	Sequence 2, Appl
83	73	5.6	718	9	US-10-117-846-2	Sequence 2, Appl
84	73	5.6	1042	10	US-09-888-615-74	Sequence 74, Appl
85	73	5.6	2472	10	US-09-815-242-5064	Sequence 5064, App
86	73	5.6	3354	9	US-10-160-758-11	Sequence 11, Appl
87	73	5.6	3354	9	US-10-160-758-12	Sequence 12, Appl
88	72.5	5.5	149	10	US-10-092-154-517	Sequence 517, App
89	72.5	5.5	149	10	US-09-764-847-517	Sequence 517, App
90	72.5	5.5	296	10	US-09-801-368-386	Sequence 386, App
91	72.5	5.5	403	10	US-09-802-213-5	Sequence 5, Appl
92	72.5	5.5	439	10	US-09-775-117A-2	Sequence 2, Appl

Qy	4	LATS	LP----	EGIMVKFTEDRMOLFSALIKGPTTPVEDGLYLFIDQLQNPYVNPVPHHC	59
Db	53	LSTS	LSLSSSSVVFCDERLLDKMVAITIGADITPANGCFDFYDFQDQSPSP-LV	591	
Qy	60	YLSQ	SGG----	RLNPVLNDKNGKVSLLGTWIGKGTERTKSK--SLLIQVLISIGGLILVINE	115
Db	592	NLETT	GHSVRNPVLNDKGVCLSTLNTWIGRPEEKNNQPTSSFLQVLVSQSLIVAE	651	
Qy	116	PYNE	AGTSDGLOGFYENSRCYN-FMALIRVVOSEMTOLVRPHPVFEQEI	ROHSTGG	174
Db	652	PYNE	PYGRSRGTPTSGTSREYDGNIRQATVKAAMLEQIRNPSPCKEVIHKEFLKR	711	

```

RESULT 2
US-09-989-920-204
; Sequence 204, Application US/09989920
; Patent No. US2002017957A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sel-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific C
; FILE REFERENCE: DEX-0291

```

```

? FILE REFERENCE: DEX-0291
? CURRENT APPLICATION NUMBER: US/09/989-920
? CURRENT FILING DATE: 2001-11-21
? PRIOR APPLICATION NUMBER: 60/252,500
? PRIOR FILING DATE: 2000-11-22
? NUMBER OF SEQ ID NOS: 284
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 204
? LENGTH: 252
? TYPE: PRT
? ORGANISM: Homo sapien
? US-09-989-920-204
Query Match %0.0% Score 248: EB 9: Length 252:

```

Host Local Similarity 43.7%; Pred. No. 3,10-16;
 Matches 59; Conservative 16; Mismatches 50; Indels 10; Gaps 5;
 QY 9 PEGIMVKTFEDRMOLFSA:KCPTRIPYEGGLYFDIQLPNIYPAVPPHPCYLSQCSG-- 66
 Db 8 PEGIMVKTFEDRMOLFSA:KCPTRIPYEGGLYFDIQLPNIYPAVPPHPCYLSQCSG-- 66
 QY 67 KLNPNLYDNGKVCVSLGTWICKGTERW:SKSSLLQVLSIQGLIIVNPPYNEAGDSQ 126
 Db 69 RNPNFYNGKVCVSLGTWICKGTERW:SKSSLLQVLSIQGLIIVNPPYNEAGDSQ 126
 QY 127 RGLQGYNSRCYNE 141
 Db 124 R-HPG--DSKNEY 134

RESULT 3
 US-09-925-301-979
 : Sequence 979, Application US/09925301
 : Patent No. US20020052308A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 : FILE REFERENCE: PA106
 : CURRENT APPLICATION NUMBER: US/09925301
 : CURRENT FILING DATE: 2001-08-10
 : PRIOR APPLICATION NUMBER: PCT/US02/05882
 : PRIOR FILING DATE: 2000-03-08
 : PRIOR APPLICATION NUMBER: 60/124,270
 : PRIOR FILING DATE: 1999-03-12
 : NUMBER OF SEQ ID NOS: 1694
 : SOFTWARE: Patent Ver. 2.0
 : SEQ ID NO 979
 : LENGTH: 283
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-925-301-979

Query Match 12.9%; Score 169; DB 10; Length 283;
 Best Local Similarity 24.5%; Pred. No. 1,1e-06;
 Matches 60; Conservative 36; Mismatches 73; Indels 76; Gaps 10;
 QY 4 LATSIPGIMVKTFEDRMOLFSA:KCPTRIPYEGGLYFDIQLPNIYPAVPPHPCYLSQ 63
 Db 84 LATSIPGIMVKTFEDRMOLFSA:KCPTRIPYEGGLYFDIQLPNIYPAVPPHPCYLSQ 63
 QY 64 CSGRNPNLYDNGKVCVSLGTWICKGTERW:SKSSLLQVLSIQGLIIVNPPYNEAG 123
 Db 144 CSGRNPNLYDNGKVCVSLGTWICKGTERW:SKSSLLQVLSIQGLIIVNPPYNEAG 123
 QY 124 DSURG--IQGYNSRCYNEALIRVQSMQTLVRPFFVEQEIRQHFTSGWLVNRI 181
 Db 191 NEAGRLLENYEE---YAARAL-----LTEL-----HGGAGG----- 221
 QY 182 ESWLEIHALIKAGALPNGVKASSFPFPAVAELSDSQQEPDGGPAGGASQSDSE 241
 Db 222 ESWLEIHALIKAGALPNGVKASSFPFPAVAELSDSQQEPDGGPAGGASQSDSE 241
 QY 242 GSAAG 246
 Db 250 GSAAG 254

RESULT 4
 US-10-157-669-33
 : Sequence 33, Application US/10157669
 : Publication No. US20030054385A1
 : GENERAL INFORMATION:
 : APPLICANT: Lal, Preeti G.
 : APPLICANT: Jackson, Jennifer L.
 : APPLICANT: Corley, Neil C.
 : TITLE OF INVENTION: HUMAN UBIQUITIN-CONJUGATING ENZYMES
 : FILE REFERENCE: PP-0356-2 CIP

: CURRENT APPLICATION NUMBER: US/10157,669
 : CURRENT FILING DATE: 2002-05-28
 : PRIOR APPLICATION NUMBER: 09/520,076; 09/359,967; 08/965,689; 08/933,750
 : PRIOR FILING DATE: 2000-03-07; 1999-07-22; 1997-11-06; 1997-09-23
 : NUMBER OF SEQ ID NOS: 36
 : SOFTWARE: PERL Program
 : SEQ ID NO 33
 : LENGTH: 165
 : TYPE: PRT
 : ORGANISM: Saccharomyces cerevisiae
 : PUBLICATION INFORMATION:
 : DATABASE ACCESSION NUMBER: Genbank ID No. US20030054385A1 g4257
 US-10-157-669-33

Query Match 12.6%; Score 165.5; DB 9; Length 165;
 Best Local Similarity 31.2%; Pred. No. 1,7e-08;
 Matches 39; Conservative 28; Mismatches 45; Indels 13; Gaps 5;
 QY 9 PEGIMVKTFEDRMOLFSA:KCPTRIPYEGGLYFDIQLPNIYPAVPPHPCYLSQCSG 67
 Db 21 PEGIMVKTFEDRMOLFSA:KCPTRIPYEGGLYFDIQLPNIYPAVPPHPCYLSQCSG 67
 QY 68 LNPNIYDNGKVCVSLGT-----WICKGTERWTSKSSLLQVLSIQGLIIVNPPYNEA 121
 Db 78 LNPNIYDNGKVCVSLGT-----WICKGTERWTSKSSLLQVLSIQGLIIVNPPYNEA 121
 QY 122 GFDSD 126
 Db 135 CANID 139

RESULT 5
 US-09-902-941-1904
 : Sequence 1904, Application US/09902941
 : Patent No. US20020172952A1
 : GENERAL INFORMATION:
 : APPLICANT: Henderson, Robert A.
 : APPLICANT: Wang, longtong
 : APPLICANT: Watanabe, Yoshihiro
 : APPLICANT: Johnson, Jeffrey C.
 : APPLICANT: Retter, Marc W.
 : APPLICANT: Marnerakis, Margarita
 : APPLICANT: Carter, Darlick
 : APPLICANT: Fanger, Gary R.
 : APPLICANT: Vedvick, Thomas S.
 : APPLICANT: Bangur, Chaitanya S.
 : APPLICANT: McNabb, Andria
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IHE THERAPY
 : TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 : FILE REFERENCE: 210121.478C17
 : CURRENT APPLICATION NUMBER: US/09902,941
 : CURRENT FILING DATE: 2001-07-10
 : NUMBER OF SEQ ID NOS: 2002
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1904
 : LENGTH: 197
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-902-941-1904

Query Match 12.5%; Score 164; DB 9; Length 197;
 Best Local Similarity 28.3%; Pred. No. 3e-08;
 Matches 49; Conservative 33; Mismatches 75; Indels 16; Gaps 6;
 QY 3 LLATSLPEGIMVKTFEDRMOLFSA:KCPTRIPYEGGLYFDIQLPNIYPAVPPHPCYLS 62
 Db 13 LLATSLPEGIMVKTFEDRMOLFSA:KCPTRIPYEGGLYFDIQLPNIYPAVPPHPCYLS 62
 QY 63 QCSGRNPNLYDNGKVCVSLGTWICKGTERWTSKSSLLQVLSIQGLIIVNPPYNEAG 122
 Db 73 QCSGRNPNLYDNGKVCVSLGTWICKGTERWTSKSSLLQVLSIQGLIIVNPPYNEAG 122
 QY 123 PDSDEGLIQGYNSRCYNEALIRVQSMQTL--VRPPPVFFQEIQRHFTSGG 174

```
US-10-017-754-1904

Query Match          12.5%; Score 164; Dh 9; Length 197;
Best Local Similarity   28.3%; Pred. No. 3e-08;
Matches      49; Conservative    33; Mismatches     75; Indels    16; Gaps         ;
```

Oy 3 LATSLEPGIMVKVTFEDRMDFESALIKGPTTPYVEDGLYLFDLQILPNIYPVPHPHICYLS 62
 :||| ||| :||| :||| :||| :||| :||| :
Db 13 MLATPPIGTICWODKIQMDUDDLAQLLGANIPTEKGVGFKLEVITPERYPPEHQNFLLT 72
 :||| :||| :||| :||| :||| :||| :

Oy 63 QCSGRINPINXDNGKVCVSILLCTICKGFIRWTSSKSLLQVLTSIOGIIVNRIYYNEAG 122
 :||| :||| :||| :||| :||| :||| :||| :
Db 73 PI---VHENIDSAGRIGLDVI-KLPKA--WRPSLNIAVTITSIQ--LLMSEP----- 118

Oy 123 FSDRRLGLEGYENSCKNSNMALIRVVQSMTOL-VRRPPVFVEQEIRQHFSTGG 174
 : | | : | | : | | : | | : | | : | | : | | :
Db 119 -RPDOPPLMADTSSFKNFYKNPAFLKNARQMITEKHAKQKADREEMLCNLFEAG 170

RESULT 8
US-10-180-927-2
; Sequence 2, Application US/10180927
; Publication No. US20030046721A1
; GENERAL INFORMATION:
; APPLICANT: Leviten, Michael W.
; TITLE OF INVENTION: HSPC150-LIKE GENE DISRUPTIONS;
; FILE REFERENCE: R-477
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATED THERE TO

```

; CURRENT APPLICATION NUMBER: US/10/180,927
;
; CURRENT FILING DATE: 2002-06-25
;
; PRIOR APPLICATION NUMBER: US 60/301,281
;
; PRIOR FILING DATE: 2001-06-26
;
; NUMBER OF SEQ ID NOS: 4
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 2
;
; LENGTH: 204
;
; TYPE: PRI
;
; ORGANISM: Mus musculus
;
US-10-180-927-2

```

[illegible]


```

; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: US/520,076; 09/359,967; 08/965,689; 08/933,750
; PRIOR FILING DATE: 2000-03-07; 1999-07-22; 1997-11-06; 1997-09-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030054385A1 2456290CB1
US-10-157-669-17

Query Match          10.5%; Score 137.5; DB 9; Length 165;
Best Local Similarity 29.2%; Pred. No. 8.7e-06;
Matches 38; Conservative 27; Mismatches 52; Indels 13; Gaps 5;

QY      4 LATSPLPGTMVTEFEDRMILFSALIKGPTPYPMGLIYFDIQLENIYPAVPPHCYLS 62
       : : |||:: : : |||:: : : |||:: : : |||:: : : |||:: : : |||::
Db      5 LTNPPGGIVAGPMNEEFMEALMGPDTCFEFGVEPAILSFPLDYLPSPKMRFTC 75
       : : |||:: : : |||:: : : |||:: : : |||:: : : |||:: : : |||::

QY     63 QCSGRLNPNLYONGKVCVSLGT-----WICKGTERWTSSKLQVLISIOGLILYNRP 116
       : :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||::
Db     76 E---MHPNTPYDGRVCISILHAPDDPMGYESSAERWSPVQSVEKILISVVSMCA--EP 130
       : :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||::

QY    117 YYNEAGFSDS 126
       : :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||::
Db    131 -NDESGANVD 139

RESULT 15
US-10-108-605-331
; Sequence 331, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENC
; FILE REFERENCE: 3113BR
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 331
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-331

Query Match          10.2%; Score 134; DB 9; Length 147;
Best Local Similarity 27.9%; Pred. No. 1.6e-05;
Matches 36; Conservative 24; Mismatches 43; Indels 26; Gaps 6;

QY    22 DLFL-SALIKGPTPYPMGLIYFDIQLENIYPAVPPHCYLSOCSRLNPMLNYNGKVC 79
       : :|||:: :|||:: :|||:: :|||:: :|||:: :|||:: :|||:: :|||::
Db    29 DLFWQATIGMPDPSYQGVEFFLIHFETDPDFPKPVATTIRI---YEPINSNSSIC 85
       : :|||:: :|||:: :|||:: :|||:: :|||:: :|||:: :|||:: :|||::

QY    80 VSLLOTWIGKTERWTSSKLQVLISIOGLIL---VNRPYYNFAG--FDSORGIQEYE 134
       : :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||::
Db    86 LDIL-----RSQSPALTISKVLLSICSLCDPNPDPLVPEIAFYKTDR----- 131
       : :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||::

QY   135 NSRCYNENKA 143
       : :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||:: :||::
Db   132 --FKYNELA 138

```

86 VYFSGTICLSILNE-----DQDWRPAITLKQIVIGVQDILJUSPNPNSPAQE:PAWRPSKN 140

QY 120 EAGHSDRGCIQ 130
 : : : : :
 Db 141 KAEYDKKVLIIQ 151

RESULT 20

US-09-785-671 2
 : Sequence 2, Application US/09785671

: Patent No. US2002028189A1

: GENERAL INFORMATION:

: APPLICANT: Kbo. Cicon-Joo

: APPLICANT: Lee, Mu En

: APPLICANT: Haber, Edgar

: TITLE OF INVENTION: URICEMINATION OF THE TRANSCRIPTION FACTOR F2A

: FILE REFERENCE: 05433-035002

: CURRENT APPLICATION NUMBER: US/09/785,671

: PRIOR FILING DATE: 2001-09-14

: PRIOR APPLICATION NUMBER: US 08/986,605

: PRIOR FILING DATE: 1997-12-08

: PRIOR APPLICATION NUMBER: US 08/425,476

: PRIOR FILING DATE: 1997-03-28

: PRIOR APPLICATION NUMBER: US 60/014,388

: PRIOR FILING DATE: 1996-03-28

: NUMBER OF SEQ ID NOS: 7

: SOFTWARE: FASTSEQ for Windows Version 4.0

: SEQ ID NO 2

: LENGTH: 158

: TYPE: PKI

: ORGANISM: Rattus rattus

US-09-785 671-2

Query Match 9.0% Score 113; DB 10; Length 158;

Best local Similarity 24.9% Pred. No. 0.00063;

Matches 33; Conservative 21; Mismatches 38; Indels 22; Gaps 5;

QY 6 ISLPEGIWKTEDRMDFSAIIGKTRTPYEDGLYFDIQLPNYPAVPPHFCYLSQCS 65

Db 29 TRNPDSTM-----NLNWECAIPGKGGIPLKRLMFKDDYPSPP-----KCK 76

QY 66 GR---INPNLYDNCKYCVSLIGTWICKGTRWTSKSSLLQVLISIQGLIIVNEP 116

Db 77 FEPLPHNPVPSCTVCLSL-----EEDKDRPAITIKOILIGIQELI--NEP 123

RESULT 21

US-09-925-302-525

: Sequence 525, Application US/09925302

: Patent No. US20020044941A.

: GENERAL INFORMATION:

: APPLICANT: Rosen et al.

: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

: FILE REFERENCE: PA104

: CURRENT APPLICATION NUMBER: US/09/925,302

: PRIOR FILING DATE: 2001-08-10

: PRIOR APPLICATION NUMBER: PCT/US00/05918

: PRIOR FILING DATE: 2000-03-08

: PRIOR APPLICATION NUMBER: 60/124,270

: PRIOR FILING DATE: 1999-03-12

: NUMBER OF SEQ ID NOS: 896

: SOFTWARE: PatentIn Ver. 2.0

: SEQ ID NO 525

: LENGTH: 205

: TYPE: PRT

: ORGANISM: Homo sapiens

US-09-925-302 525

Query Match 9.0% Score 118; DB 10; Length 205;

Best local Similarity 28.9% Pred. No. 0.00091;

Matches 33; Conservative 21; Mismatches 38; Indels 22; Gaps 5;

QY 6 ISLPEGIWKTEDRMDFSAIIGKTRTPYEDGLYFDIQLPNYPAVPPHFCYLSQCS 65

Db 77 FEPLPHNPVPSCTVCLSL-----EEDKDRPAITIKOILIGIQELI--NEP 123

Db 76 TRNPDSTM-----NLNWECAIPGKGGIPLKRLMFKDDYPSPP-----KCK 123

QY 66 GR---INPNLYDNCKYCVSLIGTWICKGTRWTSKSSLLQVLISIQGLIIVNEP 116

Db 124 FEPLPHNPVPSCTVCLSL-----EEDKDRPAITIKOILIGIQELI--NEP 170

RESULT 22

US-09-976-165-22

: Sequence 22, Application US/09976165

: Patent No. US20020107383A1

: GENERAL INFORMATION:

: APPLICANT: FUJIMURA, TSUTOMU

: APPLICANT: WATANABE, TAKESHI

: APPLICANT: HORIE, MASATO

: TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN

: FILE REFERENCE: Q-53599

: CURRENT APPLICATION NUMBER: US/09/976,165

: PRIOR FILING DATE: 2001-10-15

: PRIOR APPLICATION NUMBER: 09/565,538

: PRIOR FILING DATE: 2000-05-05

: PRIOR APPLICATION NUMBER: 09/055,699

: PRIOR FILING DATE: 1998-04-07

: PRIOR APPLICATION NUMBER: 08/820,170

: PRIOR FILING DATE: 1997-03-19

: PRIOR APPLICATION NUMBER: JP 63410/1996

: PRIOR FILING DATE: 1996-03-19

: PRIOR APPLICATION NUMBER: JP 69163/1997

: PRIOR FILING DATE: 1997-03-05

: NUMBER OF SEQ ID NOS: 95

: SOFTWARE: PatentIn Ver. 2.1

: SEQ ID NO 22

: LENGTH: 170

: TYPE: PRT

: ORGANISM: Homo sapiens

US-09-976-165-22

Query Match 8.9% Score 117; DB 10; Length 170;

Best local Similarity 29.1% Pred. No. 0.00088;

Matches 32; Conservative 20; Mismatches 46; Indels 12; Gaps 4;

QY 10 EGIWKTEDRMDF--SALIKGTRTPYEDGLYFDIQLPNYPAVPPHFCYLSQCSGR 67

Db 23 EGSAGLIDDN-DLYRWELIIGPDTLYEGGVKKAHLTFPKDYPLRPPKMKFIEI--- 78

QY 68 LNPNDYDNCKYCVSLIGT-----WIGKTRWTSKSSLLQVLISIQGLI 111

Db 79 WIPNDKNGDVCISILHEPGEDKYGYKPEERWLPPIHTVETIMISVISML 128

RESULT 23

US-10-060-019-22

: Sequence 22, Application US/10060019

: Publication No. US20030003564A1

: GENERAL INFORMATION:

: APPLICANT: Tyers, Mike

: APPLICANT: Willems, Andrew

: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING CRYPTOILIN

: FILE REFERENCE: 11757.100S01

: CURRENT APPLICATION NUMBER: US/10/060,019

: PRIOR FILING DATE: 2002-01-29

: PRIOR APPLICATION NUMBER: US/09/177,165

: PRIOR FILING DATE: 1998-10-22

: PRIOR APPLICATION NUMBER: 60/092,443

: PRIOR FILING DATE: 1998-07-10

: PRIOR APPLICATION NUMBER: 60/063,254

: PRIOR FILING DATE: 1997-10-24

: NUMBER OF SEQ ID NOS: 50

: SOFTWARE: PatentIn Ver. 2.1

: SEQ ID NO 22

: LENGTH: 295

TYPE: PRI
ORGANISM: Sacccharomyces cerevisiae
US-10-060-019-22

Query Match
Best Local Similarity 8.3%; Score 109; DB 9; Length 295;
Matches 42; Conservative 27; Mismatches 58; Indels 76; Gaps 8;

QY 36 YKGLYLFDIQLPNIYPAVPPHFCYLSQSGRLNPNLYDNGKVCVSLIGTWICKG-----90
Db 55 YKGLYKFAOMRPEDEFFSPQAFETAT--YHPNVDKGRICISLIHQ---SGDPMTD 108
QY 92 ---TERWTSKSSILQVLSISIGLIL--LVNPNYNEAGFSDRGIOGVENSRCYNEMAL 144
Db 109 EPDAETWSPQVIESVLSISVLSLEDPINSPANDVAVD-----148
QY 145 IRVQSKTOIVRRPPEVFESE--KQHFSTGGRKLVNRIESWLETHALLEKAAQALPNG--VP 202
Db 149 -----YKKNPQYKQKRVKVE-----RSKODIPKGFIMP 178
QY 203 KASSPEPPAVAEISDSGGEPE 225
Db 179 TSES-----AYISQSKLDEPE 194

RESULT 24
US-10-043-487-394
Sequence 394, Application US/10043487
Publication No. US20030055220A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
APPLICANT: PIERRE, LEGRAIN
TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
FILE REFERENCE: B4778A
CURRENT APPLICATION NUMBER: US/10/043,487
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/261,130
PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 561
SOFTWARE: PatentIn version 3.1
SEQ ID NO 394
TYPE: PRI
ORGANISM: Shigella Flexneri
US-10-043-487-394

Query Match
Best Local Similarity 8.2%; Score 107.5; DB 9; Length 153;
Matches 32; Conservative 19; Mismatches 45; Indels 17; Gaps 6;

QY 31 PRTPEVDSGLYFDIQLPNIYPAVPPHFCYLSQSGRLNPNLYDNGKVCVSLIGTWICKG 90
Db 40 QUNPYDKGAFRIETNPAEPYFPKPKTKTKI---VHPNIDKGGVCI.PVI-----S 90
QY 91 TERWTSKSSILQVLSISIGLIL--LVNPNYNEAGFSDRGIOGVENSRCYNEMAL 140
Db 91 AENKPAIKTDQVLSL--IALVNDP-QPEHPLRAD--LAEYSKDRKKFCNK 138

RESULT 25
US-10-043-487-395
Sequence 395, Application US/10043487
Publication No. US20030055220A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
APPLICANT: PIERRE, LEGRAIN
TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
FILE REFERENCE: B4778A
CURRENT APPLICATION NUMBER: US/10/043,487
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/261,130

PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 561
SOFTWARE: PatentIn version 3.1
SEQ ID NO 395
LENGTH: 153
TYPE: PRI
ORGANISM: Shigella Flexneri
US-10-043-487-395

Query Match
Best Local Similarity 7.3%; Score 96; DB 9; Length 153;
Matches 28; Conservative 25; Mismatches 49; Indels 36; Gaps 4;

QY 31 PRTPEVDSGLYFDIQLPNIYPAVPPHFCYLSQSGRLNPNLYDNGKVCVSLIGTWICKG 90
Db 41 PDOPPYHLKAFNLRAISFPPEYFPKPKTKTKI---YHPNVDKGGVCI.PVI-----S 91
QY 91 TERWTSKSSILQVLSISIGLIL--LVNPNYNEAGFSDRGIOGVENSRCYNEMALIRVWQS 150
Db 92 SENWKPCTKTCQVIFALN--VLVNRUNIREP-----L.RMD 124
QY 151 MTOLVRRPPEVEFEQEIQQ 168
Db 125 LADLITQNPFLFRKNAEE 142

RESULT 26
US-10-043-487-401
Sequence 401, Application US/10043487
Publication No. US20030055220A1
GENERAL INFORMATION:
APPLICANT: HYBRIGENICS
APPLICANT: PIERRE, LEGRAIN
TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
FILE REFERENCE: B4778A
CURRENT APPLICATION NUMBER: US/10/043,487
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/261,130
PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 561
SOFTWARE: PatentIn version 3.1
SEQ ID NO 401
TYPE: PRI
ORGANISM: Shigella Flexneri
US-10-043-487-401

Query Match
Best Local Similarity 7.3%; Score 95.5; DB 9; Length 176;
Matches 25; Conservative 23; Mismatches 38; Indels 17; Gaps 3;

QY 9 PEGIMVKTFEDRMOLFALIKGTRTPYEDGLYFDIQLPNIYPAVPPHFCYLSQSGRL 68
Db 54 PKG-----DNVYKRSITILGPGSVYGGVFFLDITFSSDYFPKPKTKTKI---Y 103
QY 69 NPMLDNGKVCVSLIGTWICKGTERWTSKSSILQVLSISIGLIL 111
Db 104 HCNINSQGVICLDIL-----KNNNSPALITISKVLLSICSL 139

RESULT 27
US-09-925-302-746
Sequence 746, Application US/09925302
Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: ROSEN ET AL.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08

```

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 746
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-746

Query Match      7.2%  Score 94.5; DB 10; Length 186;
Best Local Similarity 28.2%  Pred. No. 0.3;
Matches 29; Conservative 21; Mismatches 34; Indels 19; Gaps 6;

QY 13 MKVIFEDRWII-----+SALIKGTPTPYEDGLYFDIQ.PNIYPVPPHFCYLSQC 64
DB 36 VKVLIKREHVTIIIGC.NRFVWFFYGHQCTPYEGGVWVKVRVDE.PDKYFKSPSIGFMNKT 95

QY 65 SGRINPLNYD-NGKVCVSLIGTWIGTERWTSKSSLLQVLIISIOGLILVN----EPYIN 119
DB 96 ---FHPNIDEASGIVCLDVI-----NQTWALYDITNIFESPLQCLLAYPNP.DPLNG 145

QY 120 EAGFDSURGLGUYENSRQYNEMALIRVVQSMIOIVRRPPEVFEOLRQH 169
DB 146 DAA-----AMVJHR-----PPEYKQKKEY 165

RESULT 28
US-09-925-302-620
; Sequence 20, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09925302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05913
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 620
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (20)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCALIZATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-620

Query Match      7.2%  Score 94.5; DB 10; Length 305;
Best Local Similarity 28.2%  Pred. No. 0.3;
Matches 29; Conservative 21; Mismatches 34; Indels 19; Gaps 6;

QY 7 SLPEGMVKTPEKMLFSAIKGTPTPYEDGLYFDIQ.PNIYPVPPHFCYLSQSG 66
DB 15 SLP-----CALEENLFEHFTVRGPDSDGCGVHGVRIVLPPEYPMKPPSILLIT-ANG 69

QY 67 RLNPNIYONG-KVCVSLIG---TWIGKGTERTWTSKSS:LOVL 104
DB 70 R-----FEVGGKICLSISGHPETW----QPSNSIRTAILAI 103

RESULT 29
US-10-157-669-1
; Sequence 1, Application US/10157669
; Publication No. US20030054385A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Jackson, Jennifer L.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN URIQUITIN-CONJUGATING ENZYMS
; FILE REFERENCE: PF-0356-2 CIP
; CURRENT APPLICATION NUMBER: US/10/157,669
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 09/520,076; 09/359,967; 08/965,689; 08/933,750
; PRIOR FILING DATE: 2000-03-07; 1999-07-22; 1997-11-06; 1997-09-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERC Program
; SEQ ID NO 1
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030054385A1 001762CB1
US-10-157-669-1

Query Match      7.1%  Score 94; DB 9; Length 185;
Best Local Similarity 30.3%  Pred. No. 0.17;
Matches 30; Conservative 20; Mismatches 35; Indels 14; Gaps 6;

QY 32 TRTP-----YEDGLYFDIQ.PNIYPVPPHFCYLSQSGRL-NPNLYDNGKVCVSLIGTW 86
DB 68 TVIPDEGYGCKEQFTEVPDAYNMVPPKV---KCLTKIWHPNITETGEICISLIRH 123

QY 87 IGKGTERTWTSKSSLLQVLIISIOGLI--LVN--EPYNEA 121
DB 124 SIDGT-GWAPTRILKDVVWGLNSLFTDLLNFDDPLNIEA 161

RESULT 30
US-C9-840-787-2
; Sequence 2, Application US/09840787
; Patent No. US20020058264A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Bandman, Olga
; Shah, Purvi
; As-Young, Janice
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/840,787
; FILING DATE: 23-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/518,865
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hillings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: U937NUC1
CLONE: 1762
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-840-787-2

Query Match 7.1%; Score 94; DB 10; Length 185;
Best Local Similarity 30.3%; Pred. No. 0.17;

Matches 30; Conservative 20; Mismatches 35; Indels 14; Gaps 6;

QY 32 TRTP---YEDGILYFDQLENTIYPAVPPHFCYLSQCSGRL-NPNLYDNGKVCVSLGTM 86

DB 68 TVIHDGEGYGGCKFGFETEDVDAVMWPKV-----KCLIKIWHENITETGEICISLUREH 123

QY 87 IGKTERWISKSLQVLLISIOGLI--LVN--EPPYNEA 121

DB 124 SIGGF-GWAPTRIKDVMWGLNSJFTDLLNFDPLNTEA 161

RESULT 31

US-10-043-487-412

Sequence 412, Application US/10C43487

Publication No. US2003005220A1

GENERAL INFORMATION:

APPLICANT: HYBRIGENCS

APPLICANT: Pierre, ISGRAIN

TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptides

TITLE OF INVENTION: mammalian polypeptides

FILE REFERENCE: B4778A

CURRENT APPLICATION NUMBER: US/10C43,487

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/261,130

PRIOR FILING DATE: 2001-01-12

NUMBER OF SEQ ID NOS: 561

SOFTWARE: PatentIn version 3.1

SEQ ID NO 412

LENGTH: 156

TYPE: PRT

ORGANISM: Shigella flexneri

US 10-043-487-412

Query Match

Best Local Similarity 24.3%; Score 92.5; DB 9; Length 156;

Matches 25; Conservative 22; Mismatches 35; Indels 17; Gaps 3;

QY 9 PEGIMVKTEDRMDFSLFKGPTRTPEYDGLVFDIQLPNIXPAVPPHFCYLSQCSGRL 68

DB 34 PKG-----DN:YDWRSTILGPGSVYEGVFFLDITFSDYPPKPKVIFKTRI---Y 83

QY 69 NPNLYDNGKVCVSLGTMIGKTERWISKSLQVLLISIOGLI 111

DB 84 HCNINSQGVICLDIL-----KDNWSPALITISKVLLSTGSL 119

RESULT 32

US-09-925-301-1083

Sequence 1083, Application US/09925301

Patent No. US20020052308A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA106

CURRENT APPLICATION NUMBER: US/09/925,301

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05882

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 1694

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1083

LENGTH: 256

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (56)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-301-1083

Query Match

Best Local Similarity 24.3%; Score 92.5; DB 10; Length 256;

Matches 25; Conservative 22; Mismatches 39; Indels 17; Gaps 3;

QY 9 PEGIMVKTEDRMDFSLFKGPTRTPEYDGLVFDIQLPNIXPAVPPHFCYLSQCSGRL 68

DB 134 PKG-----DN:YDWRSTILGPGSVYEGVFFLDITFPEYPPKPKVIFKTRI---Y 183

QY 69 NPNLYDNGKVCVSLGTMIGKTERWISKSLQVLLISIOGLI 111

DB 184 HCNINSQGVICLDIL-----KDNWSPALITISKVLLSTGSL 219

RESULT 33

US-09-765-061B-72

Sequence 72, Application US/09765061B

Publication No. US20030022165A1

GENERAL INFORMATION:

APPLICANT: Board of Regents of the University of Texas System

TITLE OF INVENTION: Mutations in a No. US20030022165A1el Photoreceptor-pineal gc:

TITLE OF INVENTION: Leber congenital amaurosis (LCA4)

FILE REFERENCE: 96606/160TL

CURRENT APPLICATION NUMBER: US/09/765,061B

CURRENT FILING DATE: 2001-01-17

NUMBER OF SEQ ID NOS: 78

SOFTWARE: PatentIn version 3.1

SEQ ID NO 72

LENGTH: 384

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: PEPTIDE

LOCATION: (1) (384)

OTHER INFORMATION: Human AIP1 Protein

NAME/KEY: misc.feature

LOCATION: (322) (322)

OTHER INFORMATION: Xaa represents any of the twenty amino acids

US-09-765-061B-72

Query Match

Best Local Similarity 23.8%; Score 88.5; DB 9; Length 384;

Matches 46; Conservative 25; Mismatches 83; Indels 39; Gaps 8;

QY 73 YONGKVCVSLGTMIGKTERWISKSLQVLI-SIOGLIILVNEPYNEAGFSD----- 126

DB 201 YQEAICLNLTQKPKWQVQWLNKLEKNTILNLYCQCLLKKEEYEVLEHTSDILRH 260

QY 127 KGIQEGY-----ENSCYNE-----MALIKVQSMQIIVKRPPEVFOEIRKHFSTGG 174

DB 261 PGIVKAYVVRARAHAEVYNEAEKADQLKVELEPSMOKAVRRLLENRAE----- 314

QY 175 WRLVNRIESWLETHALLEKA-----QALPNCVFKASSPEPPA-----VAELSDSGQDEP 224

DB 315 -----KQFERILXCRNMLISQGAIQPAAPPTPEPAQSTTEPPAETAPSAEILSAGPPAEP 370

QY 225 -EDGGPAPGEASQ 236

Db 371 ATEPPSPGSHQ 383

RESULT 34

US-10-102-806-723

Sequence 723, Application US/10102806

Publication No. US2003035442A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: P103P1C1

CURRENT APPLICATION NUMBER: US/10102806

PRIOR FILING DATE: 2002-03-22

PRIOR APPLICATION NUMBER: 09/925,298

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05881

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 846

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 723

LENGTH: 190

TYPE: PRT

ORGANISM: Homo sapiens

US-10-102-806-723

Query Match 6.5% Score 85.5; DB 9; Length 190;

Best Local Similarity 24.5%; Pred. No. 1.2;

Matches 35; Conservative 25; Mismatches 52; Indels 31; Gaps 8;

Qy 16 LPEFMDLAFS-ALIKGPIRIPYEDGLYLFETQLPNTYPAVPPHPCYISQSGHL-NPNLY 73

Db 57 SFSDPDLNKLKVCIDPDKYKSGKVFSEKVGQCYPHDPKV----KCEITWYHPNID 112

Qy 74 DNCKVCSVLGTWCKGIERHTWSKSLQLVLISIQGLI!---VNEPYYNAG----- 122

Db 113 LFGNVCNLIH-----REDMKPVLTINSIYGLQYLPLEPNPDLNKEAARVLNNRR 165

Qy 123 -FDSF--RGLQRYGNS----RC 138

Db 166 LPEUNVORSWRCYICSTYFRC 188

RESULT 35

US-09-765-061B-77

Sequence 77, Application US/09765061B

Publication No. US2003022165A1

GENERAL INFORMATION:

APPLICANT: Board of Regents of the University of Texas System

TITLE OF INVENTION: Mutations in a No. US2003022165A1 el Photoreceptor-pineal gene 17

FILE REFERENCE: 96606/168TL

CURRENT APPLICATION NUMBER: US/09/765,061B

PRIOR FILING DATE: 2001-01-17

NUMBER OF SEQ ID NOS: 78

SOFTWARE: PatentIn version 3.1

SEQ ID NO 77

LENGTH: 392

TYPE: PRT

ORGANISM: Macaca mulatta

FEATURE:

NAME/KEY: PEPTIDE

LOCATION: (1)..(392)

OTHER INFORMATION: Rhesus Monkey A1P1 Protein

US-09-765 061B-77

Query Match 6.4% Score 84; DB 9; Length 392;

Best Local Similarity 21.8%; Pred. No. 4.5;

Matches 42; Conservative 26; Mismatches 53; Indels 32; Gaps 7;

Qy 73 YDNCKVCSVLGTW-CKGIERHTWSKSLQLVLISIQGLIIVNEPYYNAGFDSF----- 126

Db 201 YQFATICTLRNIQTKKQWQVOWLKLKLMINTLTINYQCCLIKKEEYEVLEHISDLRHH 260

Qy 127 RGLQRYG-----ENSRCYNE-----MALIRVQVSTQLYRRPPVFEQIRQUHFSTGG 174

Db 261 PGIVKAYIVRARAHAEYNEAEAKADLQKVLBLEPSCMKAVRRELBLENRAE----- 314

Qy 175 WRLVNRIESWLFTHALLEKAQALPNGVHPKASSSPPPPAVAHSDSQOQEPHKGGA--PG 232

Db 315 ----KQEEERLCRNMLSQCATOPPAEPAPPTAPP--AELSTGPPADPPAEPPTAPPA 368

Qy 233 EASQSDSEGGAG 245

Db 369 ELSIGPPAPPAE 381

RESULT 36

US-09-842-528-2

Sequence 2, Application US/09842528

Patent No. US20020004236A1

GENERAL INFORMATION:

APPLICANT: Meyers, Rachel A.

APPLICANT: Tsai, Fong-Ying

TITLE OF INVENTION: 27960, A NOVEL UBIQUITIN CONJUGATING

TITLE OF INVENTION: ENZYME FAMILY MEMBER AND USES THEREFOR

FILE REFERENCE: 10448-045001

CURRENT APPLICATION NUMBER: US/09/842,528

PRIOR FILING DATE: 2001-04-25

PRIOR APPLICATION NUMBER: US 60/199,500

PRIOR FILING DATE: 2000-04-25

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 151

TYPE: PRT

ORGANISM: Homo sapiens

US-09-842-528-2

Query Match 6.3% Score 83; DB 10; Length 151;

Best Local Similarity 22.9%; Pred. No. 1.5;

Matches 24; Conservative 27; Mismatches 44; Indels 10; Gaps 3;

Qy 9 PRCIMV--KTPEDRMDFSAIIXGPTPTPEYGLYIPDIQIPNIYPAVPPHPCYLSQSG 66

Db 20 PPMILNEKSVONSTIOMIVDMEGAPCTLYEGKFKFQIIFKFSKRYFPDSQVMTGE-NI 78

Qy 67 ELNPNLYDNGKVCVSLLTGTWIGKGTERTWSKSLQLVLISIQGLI 111

Db 79 PVHPHYVSNHGHCISIL-----TEDNSPALSQSVQVCLSIISML 116

RESULT 37

US-09-925-297-801

Sequence 801, Application US/09925297

Patent No. US20020081659A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: P1105

CURRENT APPLICATION NUMBER: US/09/925,297

PRIOR APPLICATION NUMBER: PCT/US00/05989

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 928

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 801

LENGTH: 161

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (3)

Query Match 6.3%; Score 82.5; DB 10; Length 434;
 Best Local Similarity 22.8%; Pred. No. 7.2;
 Matches 31; Conservative 18; Mismatches 44; Indels 43; Gaps 4;

```

QY 123 HINSLGLOEYENRSCYNEMALIRVVDSMTQLVRRPPEVFEGETHQHSTGGKRLVNRIE 182
      | | ||| : : ||| : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

? CURRENT FILING DATE: 2002-05-28
? PRIOR APPLICATION NUMBER: 09/520,076; 09/359,967; 08/965,689; 08/933,750
? PRIOR FILING DATE: 2000-03-07; 1999-07-22; 1997-11-06; 1997-09-23
? NUMBER OF SEQ ID NOS: 36
? SOFTWARE: PERL Program
? SEQ ID NO 32
? LENGTH: 180

```



```

? TYPE: PRI
? ORGANISM: Caenorhabditis elegans
? PUBLICATION INFORMATION:
? DATABASE ACCESSION NUMBER: Genbank ID No. U529030054385A1 q1628097
US-10-157-669-32

Query Match      6.18; Score 80; DB 9; Length 180;
Best local Similarity 21.6%; Pred. No. 3.7;
Matches 33; Conservative 24; Mismatches 44; Indels 52; Gaps 7;

QY 32 TRTP-----YEDGLYFDIQLPNITYPAPVPHFYLSQCSGRL-NPNLYDNGKVCVSL--- 83
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 64 IVTPQGIYRGCKKFKFTVPEPPYNNVPP-----VVKCLTKVHPNINEDGSICLSILKON 119
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 84 -----GTWIGKGIETWISKSSLLQVLISIQGLLVNEPYNEAGFSDRGLOEGYENSRC 138
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 120 SLQYG-----WRPTRNLTDVV-----HGLVSLFNDLMD 148

QY 139 YNEMALIRVVGSMVQLVRRPPEVPEQRIQHFS 171
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 149 FNDALNQAQAWWSQ---NRESFNIRVREYS 177

```

Search completed: April 10, 2003, 15:37:39
 Job time : 15.2941 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:37:10 ; Search time 18.5294 Seconds
(without alignments)
1297.052 Million cell updates/sec

Title: US-09-930-026-1

Percent score: 25%

Sequence: 1 MALLATSLPEGIWVKTEDR.....PGEASQSGDSEGAAGLAFLS 250

Scoring table: 0:100

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 20812

Minimum DB seq length: 0

Maximum DB seq length: 70

Post-processing: listing first 135 summaries

Database: 1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	2.4	22	2	C42856
2	6	2.4	28	2	cytochrome p450 1c
3	6	2.4	42	2	hypothetical prote
4	6	2.4	46	2	glucose-6-phosphat
5	6	2.4	54	2	M protein precurs
6	6	2.4	63	2	hypothetical prote
7	5	2.0	12	2	f-cell receptor be
8	5	2.0	15	2	platelet-derived g
9	5	2.0	15	2	ig heavy chain DJ
10	5	2.0	16	2	ribosomal protein
11	5	2.0	16	2	T-cell-receptor be
12	5	2.0	16	2	T-cell receptor be
13	5	2.0	18	2	T-cell receptor al
14	5	2.0	19	2	endometrial secret
15	5	2.0	19	4	probable glutamate
16	5	2.0	22	2	AlF-43 protein - h
17	5	2.0	24	2	ribosomal protein
18	5	2.0	24	2	cytochrome c552 -
19	5	2.0	30	2	proteinase inhibit
20	5	2.0	30	2	lactonohydrolase -
21	5	2.0	35	2	hypothetical prote
22	5	2.0	35	2	29k antigen PE82 -
23	5	2.0	36	2	hypothetical prote
24	5	2.0	39	2	anthranilate phosph
25	5	2.0	39	2	tax protein - stri
26	5	2.0	39	2	hypothetical prote
27	5	2.0	40	2	ribosomal protein
28	5	2.0	40	2	hypothetical 4k pr
29	5	2.0	41	2	hypothetical prote

hypothetical prote
hypothetical prote
aralkylamine N-ace
hypothetical prote
KNA-directed RNA p
gamma-zeathionin 1
mineralocorticoid
hypothetical prote
antigen WCI (simil
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
DNA-directed DNA p
light-harvesting p
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
soluble interlock
abaecin precursor
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
76K cell surface 1
hypothetical prote
protein-tyrosine k
hypothetical prote
hypothetical prote
deoxyhypusine syn
hypothetical prote
hypothetical prote
hypothetical prote
gene 58 protein -
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
cytochrome 2 - Asia
cytochrome 5 - Chin
cytochrome 1 - fore
cytochrome 1 - Moza
cytochrome 1 - cobr
cytochrome 2 - Moza
cytochrome 3 - Moza
cytochrome 10 - mon
hypothetical prote
hypothetical prote
hypothetical prote
cytochrome 1 - ring
cytochrome 2 - fore
hypothetical prote
ORF MSV203 hypothet
hypothetical prote
hypothetical prote
cytochrome homolog
hypothetical prote
hypothetical prote
hypothetical prote
vesicular membrane
Sp12 protein homol
unknown protein en
unknown protein en

```

103 5 2.0 64 2 B87006
104 5 2.0 64 2 T25319
105 5 2.0 64 2 S01103
106 5 2.0 64 2 F7581
107 5 2.0 64 2 F95056
108 5 2.0 64 2 A11676
109 5 2.0 64 2 A13288
110 5 2.0 65 2 S12405
111 5 2.0 66 2 A55545
112 5 2.0 66 2 A80076
113 5 2.0 66 2 A80618
114 5 2.0 66 2 A33469
115 5 2.0 67 2 D70929
116 5 2.0 67 2 H82963
117 5 2.0 67 2 H82755
118 5 2.0 67 2 H95273
119 5 2.0 68 2 A74016
120 5 2.0 68 2 G83633
121 5 2.0 68 2 E89843
122 5 2.0 68 2 A80427
123 5 2.0 69 2 C87574
124 5 2.0 69 2 E69385
125 5 2.0 70 2 C82962
126 4 1.6 5 2 D4817
127 4 1.6 5 2 B4817
128 4 1.6 6 2 P70593
129 4 1.6 7 2 P70667
130 4 1.6 8 2 A28004
131 4 1.6 8 2 P70522
132 4 1.6 8 2 P70639
133 4 1.6 9 2 S35538
134 4 1.6 9 2 A53797
135 4 1.6 9 2 P10326

ALIGNMENTS

RESULT 1
C42856
hypothetical protein 3 EPR-region [imported] - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C:Accession: C42856
R.Liu, Z.; Diaz, L.A.; Haas, A.L.; Giudice, G.J.
C:Ref: J. Biol. Chem. 267, 15829-15835, 1992
A:Title: cDNA cloning of a novel human ubiquitin carrier protein. An antigenic domain se
This human epidermal transcript.
A:Reference number: A42856; MUID: 92348449; PMID: 1379239
A:Accession: C42856
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-22 <HL>
A:Experimental source: keratinocyte
A>Note: sequence extracted from NCBI backbone (NCBI:109895, NCBI:109896)

Query Match 2.4%; Score 5; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 GGPAPG 232
DB 13 GGPAPG 16
|||||

RESULT 2
PX0033
cytochrome P450 testosterone 6beta-hydroxylase 2 - rat (fragment)
N:Contains: oxidoreductase (EC 1.1.1.1)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Mar-1999
C:Accession: PX0033
R.Nagata, K.; Gonzalez, F.J.; Yamazoe, Y.; Kato, R.

```

```

J. Biochem. 107, 718-725, 1990
A:Title: Purification and characterization of four catalytically active testosterone
nally related forms.
A:Reference number: PX0032; MUID: 90375438; PMID: 2398038
A:Accession: PX0033
A:Molecule type: protein
A:Residues: 1-28 <NAG>
A:Experimental source: liver, Sprague-Dawley male rat, 8 weeks old
A:Gene: CYP3A
C:Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C:Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane pr

Query Match 2.4%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LATSL 8
DB 14 LATSL 19
|||||

RESULT 3
T07474
hypothetical protein 42a - Japanese black pine chloroplast
C:Species: chloroplast Pinus thunbergiana (Japanese black pine)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 18-Aug-2000
C:Accession: T07474
R.Kakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A:Title: Loss of all ndh genes as determined by sequencing the entire chloroplast ge
A:Reference number: 216030; MUID: 95024047; PMID: 7937893
A:Accession: T07474
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-42 <WAK>
A:Cross-references: EMBL: D17510; NID: g529643; PIDN: BAA04352.1; PID: gl262635
C:Gene: ndh
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 2.4%; Score 6; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 GRNPN 71
DB 23 GRNPN 28
|||||

RESULT 4
A61078
glucose-6-phosphate isomerase (EC 5.3.1.9) - chicken (fragment)
N:Alternate names: neuroleukin; phosphoglucose isomerase; phosphohexose isomerase
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C:Accession: A61078
R.Hallböök, F.; Persson, H.; Barbany, G.; Ebendal, T.
J. Neurosci. Res. 23, 142-151, 1989
A:Title: Development and regional expression of chicken neuroleukin (glucose-6-phosp
A:Reference number: A61078; MUID: 89329085; PMID: 2754762
A:Accession: A61078
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-46 <HAL>
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: glycolysis; intramolecular oxidoreductase; isomerase; neurotrophic facto

Query Match 2.4%; Score 6; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 ETHALL 191

```

```

DE      25 EHHAL 30
|||||
RESULT 5
S60852
M protein precursor: Streptococcus pyogenes (serotype M80) (fragment)
C:Species: Streptococcus pyogenes
A:Variety: serotype M80
C>Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 26-Aug-1999
C:Accession: S60852
R:Watmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Keboe, M.A.
Mol. Microbiol. 14, 619-631, 1994
A:Title: Non congruent relationships between variation in emm gene sequences and the pop
A:Reference number: S60784; MUID:95198537; PMID:7891551
A:Accession: S60852
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-54 <WHA>
A:Cross-references: EMBL:U12004; NID:q533685; P:DN:AAA99620.; PID:q533686
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Superfamily: M5 protein

Query Match      2.4%; Score 6; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 TSKSSL 100
|||||
DB 43 TSKSSL 48

RESULT 6
D66673
Hypothetical protein ydjB (imported) - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: D66673
R:Belotin, A.; Wlacker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:1137471
A:Accession: D66673
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <STO>
A:Cross-references: GB:AEW03176; PID:q12723261; P:DN:AAK04486.1; GSPDB:GNG0146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ydjB

Query Match      2.4%; Score 6; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 RQHST 172
|||||
DB 10 RQHST 15

RESULT 7
PH1458
T-cell receptor beta chain (clone 332/1K) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 17-Mar-1999
C:Accession: PH1458; S26543
R:Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Parnetier, C.; Regnault, A.; K
J. Exp. Med. 177, 811-820, 1993
A:Title: T cell receptor selection by and recognition of two class I major histocompatib
A:Reference number: PH1430; MUID:93171821; PMID:8436911
A:Accession: PH1458
A:Molecule type: mRNA
A:Residues: 1-12 <CAS>

```

```

A:Experimental source: cytolytic T-lymphocyte, clone 332/1K
R:Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.;
J. Exp. Med. 176, 439-447, 1992
A:Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell recee;
A:Reference number: S26512; MUID:92364546; PMID:1380061
A:Accession: S26543
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-12 <CAZ>
A:Cross-references: EMBL:X67993
A:Experimental source: cytolytic T-lymphocyte, clone CW3/C44
C:Superfamily: immunoglobulin homology
C:Keywords: receptor; T-cell

```

```

Query Match      2.0%; Score 5; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 2e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 235 SGGSD 239
|||||
DB 4 SGGSD 8

```

```

RESULT 8
A22789
platelet-derived growth factor chain B - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 31-Mar-1988 #sequence_revision 02-Jun-1988 #text_change 18-Jun-1993
C:Accession: A22789
R:Stroobant, P.; Waterfield, M.D.
EMBO J. 12, 2463-2467, 1984
A:Title: Purification and properties of porcine platelet-derived growth factor.
A:Reference number: A22789
A:Accession: A22789
A:Molecule type: protein
A:Residues: 1-15 <SR>
C:Superfamily: platelet-derived growth factor
C:Keywords: growth factor; mitogen

```

```

Query Match      2.0%; Score 5; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.5e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 211 PAVAF 215
|||||
DB 5 PAVAF 9

```

```

RESULT 9
PH1314
Ig heavy chain DJ region (clone C200-98) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1314
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor l
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1314
A:Molecule type: DNA
A:Residues: 1-15 <WAS>
C:Keywords: heterotetramer; immunoglobulin

```

```

Query Match      2.0%; Score 5; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.5e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 81 SLGCT 85
|||||
DB 6 SLGCT 10

```

```

RESULT 10

```

S51057
 ribosomal protein S6 - Thermus aquaticus
 C:Species: Thermus aquaticus
 C:Date: 15-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
 C:Accession: S51057
 R:Tsuboli, P.; Herturth, E.; Choli, Z.
 Eur. J. Biochem. 226, 169-177, 1994
 A:Title: Purification and characterization of the 30S ribosomal proteins from the bacterium *Thermus aquaticus*
 A:Reference number: S51057
 A:Accession: S51057
 A:Molecule type: protein
 A:Residues: 1-16 <NTS>
 A:Note: the source is given as *Thermus thermophilus*
 C:Keywords: protein biosynthesis; ribosome

Query Match 2.0%; Score 5; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 LNPNL 72
 DB 10 LNPNL 14
 IIII

RESULT 1:
 D49655
 T-cell receptor beta chain: variable region, PCR V beta (clone 2) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
 C:Accession: D49655
 R:Grom, A.A.; Thompson, S.D.; Luyrink, L.; Passo, M.; Choi, E.; Glass, D.N.
 Proc. Natl. Acad. Sci. U.S.A. 96, 11104-11108, 1999
 A:Title: Dominant T-cell receptor beta chain variable region V beta 14+ clones in juvenile mice
 A:Reference number: D49655; MUID:94068553; PMID:9248225
 A:Accession: D49655
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-16 <GR>
 A:Experimental source: peripheral blood lymphocytes
 A:Note: sequence extracted from NCBI backbone (NCBI:P:140448)
 C:Keywords: T-cell receptor

Query Match 2.0%; Score 5; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 ASSSP 208
 DB 2 ASSSP 6
 IIII

RESULT 12
 PH0767
 T-cell receptor beta chain (J5) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
 C:Accession: PH0767
 R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
 J. Exp. Med. 174, 1371-1383, 1991
 A:Title: T cell receptor genes in a series of class I major histocompatibility complex alleles: exclusion and antigen-specific repertoire.
 A:Reference number: PH0746; MUID:92078046; PMID:1836010
 A:Accession: PH0767
 A:Molecule type: mRNA
 A:Residues: 1-16 <CAS>
 A:Cross-references: EMBL:X60561; NID:q52752; PIDN:CAA43251.1; PID:q52753
 A:Experimental source: T lymphocyte
 C:Keywords: T-cell receptor

Query Match 2.0%; Score 5; DB 2; Length 16;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 ASSSP 208
 DB 2 ASSSP 6
 IIII

RESULT 13
 C32537
 T-cell receptor alpha chain J region (HAJ117) - human
 C:Species: Homo sapiens (man)
 C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 30-May-1997
 C:Accession: C32537
 R:Kimura, N.; Iyomaga, B.; Yoshikai, Y.; Du, R.P.; Mak, T.W.
 Eur. J. Immunol. 17, 375-383, 1987
 A:Title: Sequences and repertoire of the human T cell receptor alpha- and beta-chain genes
 A:Reference number: A91263; MUID:87190670; PMID:3494611
 A:Accession: C32537
 A:Molecule type: mRNA
 A:Residues: 1-18 <KIT>
 A:Cross-references: GB:M27378
 C:Keywords: T-cell receptor

Query Match 2.0%; Score 5; DB 2; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GKQTE 92
 DB 5 GKQTE 9
 IIII

RESULT 14
 A61377
 endometrial secretory protein - sheep (fragment)
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 17-Mar-1999
 C:Accession: A61377
 R:Vallet, J.L.; Barker, P.J.; Lammung, G.E.; Skinner, N.; Huskisson, N.S.
 J. Endocrinol. 130, R1-R4, 1991
 A:Title: A low molecular weight endometrial secretory protein which is increased by oestrogen
 A:Reference number: A61377; MUID:92013712; PMID:1919188
 A:Accession: A61377
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-19 <VAL>

Query Match 2.0%; Score 5; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 PEDGG 228
 DB 15 PEDGG 19
 IIII

RESULT 15
 I37424
 Probable glutamate dehydrogenase (NAD(P)+) pseudogene 5 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 01-Dec-2000
 C:Accession: I37424; S29932
 R:Tzimagiorgis, G.; Leversha, M.A.; Chroniary, K.; Goulielmos, G.; Sargent, C.A.; Per Hum. Genet. 91, 433-438, 1993
 A:Title: Structure and expression analysis of a member of the human glutamate dehydrogenase gene family
 A:Reference number: I37424; MUID:93300479; PMID:8314555
 A:Accession: I37424
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-19 <TZI>
 A:Cross-references: EMBL:X67491; NID:g31822; PIDN:CAA47830.1; PID:g1335082
 A:Note: submitted to the EMBL Data Library, July 1992
 C:Genetics:
 A:Gene: GDB:G20DP5
 A:Cross-references: GDB:138339

A:Map position: 10p11.2-10p11.2
C:Keywords: pseudogene

Query Match 2.0%; Score 5; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 YNEAG 122
|||||
DB 11 YNEAG 15

RESULT 16

S40638

ATF-43 protein - human (fragments)

C:Species: Homo sapiens (man)

C:Date: 19-May-1994 #sequence_revision 10 Nov-1995 #text_change 10-Nov-1995

C:Accession: S40638

R:Hurst, H.C.; Jolly, N.F.; Jones, K.C.

Nucleic Acids Res. 19, 4601-4609, 1991

A:Title: Identification and functional characterisation of the cellular activating trans

A:Reference number: S40638; MUID:91367654; PMID:1653949

A:Accession: S40638

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-22 <HUR>

Query Match 2.0%; Score 5; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ATSLF 9
|||||
DB 10 ATSLP 15

RESULT 17

B48471

Ribosomal protein S6 - Thermus aquaticus (fragment)

C:Species: Thermus aquaticus

C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 21-Aug-1998

C:Accession: B48471

R:Garber, M.B.; Agolarov, S.C.; Elisavikina, I.A.; Fomenkova, N.P.; Nikonov, S.V.; Sedeln

Biochimie 74, 327-336, 1992

A:Title: Ribosomal proteins from Thermus thermophilus for structural investigations.

A:Reference number: A48401; MUID:92345325; PMID:1637860

A:Accession: B48471

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-24 <GAR>

A:Note: sequence extracted from NCB1 backbone (NCB1P:109931)

C:Superfamily: Escherichia coli ribosomal protein S6

Query Match 2.0%; Score 5; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 LNPML 72
|||||
DB 10 LNPML 14

RESULT 19

A60556

cytochrome c552 - Thiobacillus ferrooxidans (fragment)

C:Species: Thiobacillus ferrooxidans

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Jun-1993

C:Accession: A60556

R:Sato, A.; Fukumori, Y.; Yano, T.; Kat, M.; Yamataka, T.

Biochim. Biophys. Acta 976, 129-134, 1989

A:Title: Thiobacillus ferrooxidans cytochrome c-552: purification and some of its molecu

A:Reference number: A60556; MUID:90001272; PMID:2551385

A:Accession: A60556

A:Molecule type: protein

A:Residues: 1-24 <SAT>

C:Keywords: electron transfer; heme

Query Match 2.0%; Score 5; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 GGPAP 231
|||||
DB 5 GGPAP 9

RESULT 19

S24979

proteinase inhibitor 1 - potato (cultivar Russet Burbank) (fragment)

C:Species: Solanum tuberosum (potato)

A:Variety: cultivar Russet Burbank

C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 16-Jul-1999

C:Accession: S24979

R:Lee, J.S.; Yang, Y.S.

submitted to the EMBL Data Library, July 1992

A:Description: Nucleotide sequence of 5' flanking region of potato proteinase inh:bi

A:Reference number: S24979

A:Accession: S24979

A:Molecule type: DNA

A:Residues: 1-30 <LPE>

A:Cross-references: EMBL:Z14027; NID:921560; PTDN:CAA78402.1; PID:92156;

A:Experimental source: cultivar Russet Burbank

C:Genetics:

A:Introns: 17/1

C:Superfamily: eglin C

Query Match 2.0%; Score 5; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LLATS 7
|||||
DB 13 LLATS 17

RESULT 20

S29286

lactonohydrolase - fungus (Fusarium oxysporum)

C:Species: Fusarium oxysporum f.sp. cucumerinum

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Aug-1997

C:Accession: S29286

R:Shimizu, S.; Kataoka, M.; Shimizu, K.; Hirakata, M.; Sakamoto, K.; Yamada, H.

Eur. J. Biochem. 209, 383-390, 1992

A:Title: Purification and characterization of a novel lactonohydrolase, catalyzing ti

A:Reference number: S29286; MUID:93011152; PMID:1396712

A:Accession: S29286

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-30 <SHI>

Query Match 2.0%; Score 5; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 PPAVA 214
|||||
DB 23 PPAVA 27

RESULT 21

B84674

hypothetical protein At2g27540 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: B84674

R:Liu, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.

M.; Koo, H.; Mollat, K.S.; Cronin, L.A.; Shen, M.; VanNken, S.E.; Umayam, S.; Tallon, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84674
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-35 <SI>
A:Cross-references: GR:AB002093; NID:g3860276; PIDN:AAC73044.1; GSPDB:GN00139
A:Gene: At2g27540
A:Genetics:
A:Map position: 2

Query Match 2.0%; Score 5; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALLA 5
Db : MALLA 5

RESULT 22
B41161
29K antigen PE32 - *Campylobacter jejuni* (fragment)
C:Species: *Campylobacter jejuni*
C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 23-Mar-1993
C:Accession: B41161
R:Pei, Z.; Ellis, R.T.; Blaser, M.J.
J. Biol. Chem. 266, 16363-16369, 1991
A:Title: Identification, purification, and characterization of major antigenic proteins
A:Reference number: A41161; MUID:91358413; PMID:1885571
A:Accession: B41161
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-35 <PEI>

Query Match 2.0%; Score 5; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 GGPAP 231
Db 8 GGPAP 12

RESULT 23
283682
Hypothetical protein BR0260 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 07-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D83682
R:Takami, H.; Nakase, K.; Takaki, Y.; Maeno, G.; Susaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D83682
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-36 <STO>
A:Cross-references: GR:AP001507; GB:BA000004; NID:g10172612; PIDN:BA03979.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BR0260

Query Match 2.0%; Score 5; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 111 ILVNE 115
Db 3 ILVNE 7

RESULT 24
A05003
anthranilate phosphoribosyltransferase (EC 2.4.2.18) - *Erwinia carotovora* (fragment)
C:Species: *Erwinia carotovora*
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 05-Dec-1998
C:Accession: A05003
R:Largen, M.; Mills, S.E.; Rowe, J.; Yanofsky, C.
J. Biol. Chem. 253, 409-412, 1978
A:Reference number: A05003; MUID:78066891; PMID:338606
A:Accession: A05003
A:Molecule type: protein
A:Residues: 1-39 <LAR>
A:Genetics:
A:Gene: trpD
A:Complex: homodimer
C:Function:
A:Pathway: tryptophan biosynthesis
C:Superfamily: trpG-trpD bifunctional enzyme; trpD homology; trpG homology
C:Keywords: glycosyltransferase; homodimer; pentosyltransferase; tryptophan biosynth

Query Match 2.0%; Score 5; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 QSMTO 153
Db 24 QSMTO 28

RESULT 25
S77904
tax protein - simian immunodeficiency virus SIVagm (fragment)
C:Species: simian immunodeficiency virus SIVagm
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 20-Sep-1999
C:Accession: S77904
R:Vandamme, A.M.
submitted to the EMRL Data Library, May 1994
A:Reference number: S44288
A:Accession: S77904
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-39 <VAN>
A:Cross-references: EMBL:232851; NID:g483445; PIDN:CAA83684.1; PMID:9483446
A:Experimental source: cell-line pp 1664
A:Note: the source is designated as primate T-cell lymphotropic virus
C:Superfamily: leukemia virus trans-activating transcription regulator

Query Match 2.0%; Score 5; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ALLAT 6
Db 25 ALLAT 29

RESULT 26
G82619
hypothetical protein XF1938 [imported] - *Xylella fastidiosa* (strain 9a5c)
C:Species: *Xylella fastidiosa*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G82619
R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A:Reference number: A82615; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: G82619
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-49 <SIW>
A:Cross-references: GB:AE004013; GB:AE004849; NID:q9107030; PIDN:AAF84740.1; GSPDB:GN001
A:Experimental source: Strain 9450
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Brites, M.R.S.; Bueno, M.R.P.; Canargo, A.A.; Canargo, L.H.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Decena, C.; El-Deiry, H.; Facicani, A.P.; Ferreira, A.J.S. submitted to GeneBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraqa, J.S.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kogler, E.; Kitajima, J.P.; Krieger, J.F.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B.; Authors: Martins, R.M.F.; Matsukuma, A.Y.; Morok, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.J.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tshako, M.H.; Vailhada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:References: number: A59228
A:References: annotation:
C:Genetics:
A:Gene: XP1338
Query Match 2.0%; Score 5; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 27 LIKGP 31
DB 10 LIKGP 14
RESULT 27
S43028
ribosomal protein S3 - petunia mitochondrion (fragment)
C:Species: mitochondrion Petunia sp. (petunia)
C:Date: 07-Sep-1994 #sequence_revision 12-Dec-1997 #text_change 13-Aug-1999
C:Accession: S43028; S11632
R:Conklin, P.L.; Hauson, M.R.
Curr. Genet. 23, 477-482, 1993
A:Title: A truncated recombination repeat in the mitochondrial genome of a Petunia CMS 1
A:Reference number: S43028; MUID:93306754; PMID:8319306
A:Accession: S43028
A>Status: translation: not shown
A:Molecule type: DNA
A:Residues: 1-40 <CON>
A:Cross-references: EMBL:X67027; NID:q14199; PIDN:CAA47419.1; PID:q14201
C:Genetics:
A:Gene: rps3
A:Genome: mitochondrion
C:Superfamily: Escherichia coli ribosomal protein S3
C:Keywords: mitochondrion; protein biosynthesis; ribosome
Query Match 2.0%; Score 5; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 103 VLISI 107
DB 25 VLISI 29
RESULT 28
C32338
hypothetical 4K protein - phage T4
C:Species: phage T4
C:Date: 31-Jul-1989
C:Accession: C32338
R:Gauss, P.; Gayle, M.; Winter, R.B.; Gold, L.
Mol. Gen. Genet. 206, 24-34, 1987
A:Title: The bacteriophage T4 dexA gene: sequence and analysis of a gene conditionally
A:Reference number: A32338; MUID:87231086; PMID:3553862
C:Accession: C32338
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-40 <SAU>

A:Cross-references: GB:X04834; NID:q15253; PIDN:CAA28537.1; PID:q15256

Query Match 2.0%; Score 5; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 SKSSL 100

DB 32 SKSSL 36

RESULT 29

D55539

hypothetical protein (cars 5' region) - Azospirillum brasilense (fragment)

C:Species: Azospirillum brasilense

C:Date: 13-Sep-1995 #sequence_revision 13-Sep-1995 #text_change 13-Sep-1995

C:Accession: D55539

R:Chattopadhyay, S.; Mukherjee, A.; Ghosh, S.

J. Bacteriol. 176, 7484-7490, 1994

A:Title: Molecular cloning and sequencing of an operon, cars of Azospirillum brasil or global control of carbohydrate catabolism

A:Reference number: A55539; MUID:95095916; PMID:8002571

A:Accession: D55539

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-41 <CHA>

Query Match 2.0%; Score 5; DB 2; Length 41;

Best Local Similarity 100.0%; Pred. No. 6.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALLAT 6

DB 9 ALLAT 12

RESULT 30

F72367

hypothetical protein - Thermotoga maritima (strain MS88)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: F72367

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; H Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: F72367

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-42 <ARN>

A:Cross-references: GB:AE001727; GB:AE000512; NID:q4981015; PIDN:AAD35589.1; PID:q49

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM0504

Query Match 2.0%; Score 5; DB 2; Length 42;

Best Local Similarity 100.0%; Pred. No. 7.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 EDGSP 229

DB 11 HDGSP 15

RESULT 31

D71343

hypothetical protein TP0281 - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999

C:Accession: D71343

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.;

son, J.; Khalak, H.; Richardson, D.; Howell, R.K.; Chidambaram, M.; Otterback, T.; McQuay, L.; Weidman, J.; Smith, H.G.; Venter, J.C.
Science 287, 375-385, 1998

A:Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.

A:Reference number: A91250; MUID:98332770; PMID:9665876

A:Accession: S71343

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-42 <COL>

A:Cross-references: CB:AF001209; GB:AE000520; NID:g3322547; PIDN:AA065278.1; PID:g332256

A:Experimental source: Strain Nichols

C:Genetics:

A:Gene: TP028

Query Match 2.0% Score 5; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 7.1e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ELSDS 229

DB 30 MDS 34

|||||

RESULT 32

S58006

A:Title: N-acetyltransferase (EC 2.3.1.87) - fruit fly (*Drosophila melanogaster*) (fruit fly)

C:Species: *Drosophila melanogaster*

C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 05-May-2000

C:Accession: S68306

R:Intermed, E.; Jenoe, P.; Meyer, U.A.

FEBS Lett. 375, 148-150, 1995

A:Title: Isolation and characterization of an avian, cytoplasmic N-acetyltransferase from *D.*

A:Reference number: S58006; MUID:96087104; PMID:7498465

A:Accession: S68306

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-45; 9,10,23,24-43 <HIN>

A:Note: 41-Glu was also found

C:Keywords: acyltransferase; coenzyme A

Query Match 2.0% Score 5; DB 2; Length 43;

Best Local Similarity 100.0%; Pred. No. 7.3e-02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 PDSG 228

DB 37 PDSG 41

|||||

RESULT 33

H97484

A:Title: Hypothetical protein ACR_C_1877 (imported), - *Agrobacterium tumefaciens* (strain C58, *Corynebacterium*)

C:Species: *Agrobacterium tumefaciens*

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: B57484

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*

A:Reference number: A97359; PMID:11743194

A:Accession: B57484

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-45 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK86827.1; PID:g15156037; GSMDH:GN00169

C:Genetics:

A:Gene: ACR_C_1877

A:Map position: circular chromosome

Query Match 2.0% Score 5; DB 2; Length 45;

Best Local Similarity 100.0%; Pred. No. 7.6e-02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 PNIYP 52

DB 13 PNIYP 17

|||||

RESULT 34

S14000

A:Title: RNA-directed RNA polymerase (EC 2.7.7.48) beta chain - phage SP (fragment)

N:Alternate names: replicase beta chain

C:Species: phage SP

C:Date: 18-Feb-1994 #sequence_revision 24-Apr-1998 #text_change 24-Oct-1998

C:Accession: S14000

R:Inokuchi, Y.; Hirashima, A.; Watanabe, I.

J. Mol. Biol. 158, 711-730, 1982

A:Title: Comparison of the nucleotide sequences at the 3'-terminal region of RNAs for

A:Reference number: S07250; MUID:83010313; PMID:7120417

A:Accession: S14000

A:Status: translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-46 <INO>

A:Cross-references: EMBL:J02500; NID:g215790

C:Superfamily: phage MS2 RNA-directed RNA polymerase

C:Keywords: nucleotidyltransferase

Query Match 2.0% Score 5; DB 2; Length 46;

Best Local Similarity 100.0%; Pred. No. 7.8e-02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LATSU 8

DB 13 LATSU 17

|||||

RESULT 35

A58319

A:Title: gamma-zeathionin 1 - maize

C:Species: Zea mays (maize)

C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 03-Jun-2000

C:Accession: A58319

R:Castro, M.S.; Fontes, W.; Morhy, L.; Bloch Jr., C.

Protein Lett. 3, 267-274, 1996

A:Title: Complete amino acid sequences of two gamma-thionins from maize (*Zea mays* L.)

A:Reference number: A58319

A:Accession: A58319

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-47 <CAS>

A:Experimental source: seed

C:Superfamily: gamma-thionin

Query Match 2.0% Score 5; DB 2; Length 47;

Best Local Similarity 100.0%; Pred. No. 7.9e-02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 LOEGY 133

DB 25 LOEGY 29

|||||

RESULT 36

I53270

A:Title: mineralocorticoid receptor - rat (fragment)

C:Species: Rattus sp. (rat)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 31-Oct-1997

C:Accession: I53270

R:Kwak, S.P.; Patel, P.D.; Thompson, R.C.; Akli, H.; Watson, S.J.

Endocrinology 133, 2344-2350, 1993

A:Title: 5'-Heterogeneity of the mineralocorticoid receptor messenger ribonucleic ac.

A:Reference number: I53270; MUID:94008805; PMID:8404687

A:Accession: I53270

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-47 <RES>

A:Cross-references: G5:S66250; NID:g433297
 C:Genetics:
 A:Gene: gumA^{MR}
 C:Superfamily: unassigned erhA-related proteins; erhA transforming protein homology

Query Match 2.0%; Score 5; DB 2; Length 47;
 Best Local Similarity 100.0%; Pred. No. 7.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SLEGG 11
 |||||
 DB 8 SLEGG 12

RESULT 37
 D28667
 hypothetical protein K - Streptococcus pneumoniae (fragment)
 C:Species: Streptococcus pneumoniae
 C:Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 15-Oct-1999
 C:Accession: D28667
 R:Friebel, S.D.; Hadi, S.M.; Greenberg, B.; Lacks, S.A.
 J. Bacteriol. 170, 190-196, 1988
 A:Title: Nucleotide sequence of the hexA gene for DNA mismatch repair in Streptococcus pneumoniae
 A:Reference number: A51871; MUID:88086867; PMID:3275508
 A:Accession: D28667
 A:Molecule type: DNA
 A:Residues: 1-47 <ORF>
 A:Cross-references: GB:M18729; NID:q153654; PIDN:AAA88598.1; PTD:q1196935
 C:Genetics:
 A:Start Codon: TTG

Query Match 2.0%; Score 5; DB 2; Length 47;
 Best Local Similarity 100.0%; Pred. No. 7.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 SSIHQ 1c2
 |||||
 DB 38 SSIHQ 42

RESULT 38
 I47015
 antigen WCL (similarity) - sheep (fragment)
 N:Alternate names: TI9 protein
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 19-Jan-2001
 C:Accession: I47015
 R:O'Keefe, M.A.; Metcalfe, S.A.; Glew, M.D.; Bowden, I.; McInnes, S.; Kimpton, W.G.; Cawston, T.; Lymph node homing cells biologically enriched for gamma delta T cells express an invariant TCR alpha chain
 A:Reference number: I47014; MUID:95169648; PMID:7865462
 A:Accession: I47015
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-47 <ORF>
 A:Cross-references: GB:S76317; NID:g913188
 C:Genetics:
 A:Gene: TI9
 A:Introns: 10/1

Query Match 2.0%; Score 5; DB 2; Length 47;
 Best Local Similarity 100.0%; Pred. No. 7.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 ASQGS 238
 |||||
 DB 34 ASQGS 38

RESULT 39
 H70241
 hypothetical protein: BHL7 - Lyme disease spirochete plasmid 1/1p28.4
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
 C:Accession: H70241
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Kison, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Venter, A.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, E. Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
 A:Reference number: A70100; MUID:98065943; PMID:943665
 A:Accession: H70241
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-48 <KLE>
 A:Cross-references: GB:AF000789; NID:g2690079; PLEN:AAC65204.1; PID:g2690101; TIGR:H70241
 A:Experimental source: strain B31
 C:Genetics:
 A:Genome: plasmid

Query Match 2.0%; Score 5; DB 2; Length 48;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 KSSLL 101
 |||||
 DB 44 KSSLL 48

RESULT 40

E64618

hypothetical protein HP0789 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: E64618

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, P.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKnight, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, J.C.
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: E64618

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-48 <ORF>

A:Cross-references: GB:AE000591; GB:AE000511; NID:g2313918; PIDN:AA007847.1; PID:g2313918

Query Match 2.0%; Score 5; DB 2; Length 48;

Best Local Similarity 100.0%; Pred. No. 8.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 ALLEK 193

|||||

DB 40 ALLEK 44

Search completed: April 10, 2003, 10:42:16

Job time : 24.5294 secs

SendCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 13:39:30 ; Search time 10 Seconds
(without alignments)
1528.401 Million cell updates/sec

Title: US-09-930-026-1
Perfect score: 250
Sequence: 2 MALLATSLPGEIHWKTFEDR.....PSEASGSDSERRGAAGLAFS 250

Scoring table: OLIGO
Gapop 60.0 / Gapext 60.0

Searched: 248812 seqs, 61136040 residues

Word size: 0

Total number of hits satisfying chosen parameters: 86914

Minimum DB seq length: 0

Maximum DB seq length: 70

Post-processing: listing first 135 summaries

Database: Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCU7_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCU5_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	2.4	9	9	US-09-809-638-44
2	6	2.4	10	9	US-09-809-638-99
3	6	2.4	16	10	US-09-929-924-12
4	6	2.4	25	10	US-09-803-165-27
5	6	2.4	29	10	US-09-864-761-49563
6	6	2.4	32	10	US-09-142-755-9
7	6	2.4	35	9	US-09-843-676-176
8	6	2.4	35	9	US-09-438-486-178
9	6	2.4	35	9	US-10-053-758-178
10	6	2.4	35	9	US-10-054-295-178
11	6	2.4	35	9	US-10-054-611-178
12	6	2.4	39	10	US-09-864-761-46227
13	6	2.4	61	9	US-09-989-919-101
14	6	2.4	61	10	US-09-864-761-44746
15	6	2.4	61	10	US-09-815-242-13075
16	6	2.4	63	10	US-09-864-761-42526
17	6	2.4	65	10	US-09-864-761-42636
18	6	2.4	69	9	US-10-092-154-629
19	6	2.4	69	10	US-09-764-847-629

5	2.0	5	10	US-09-799-118-8	Sequence 8, Appl
5	2.0	5	10	US-09-859-214-41	Sequence 41, Appl
5	2.0	7	10	US-09-989-789-213	Sequence 213, Appl
5	2.0	7	10	US-09-989-789-214	Sequence 214, Appl
5	2.0	7	10	US-09-989-789-222	Sequence 222, Appl
5	2.0	8	9	US-09-945-917-16	Sequence 16, Appl
5	2.0	9	9	US-10-125-635A-93	Sequence 93, Appl
5	2.0	9	9	US-09-945-917-33	Sequence 33, Appl
5	2.0	9	9	US-09-809-638-645	Sequence 645, Appl
5	2.0	9	9	US-09-956-288-185	Sequence 185, Appl
5	2.0	10	10	US-09-884-260A-1	Sequence 1, Appl
5	2.0	11	9	US-09-945-927-15	Sequence 15, Appl
5	2.0	11	9	US-10-044-955-22	Sequence 22, Appl
5	2.0	11	10	US-09-941-611-22	Sequence 22, Appl
5	2.0	12	9	US-10-125-635A-110	Sequence 310, Appl
5	2.0	13	9	US-09-945-917-8	Sequence 8, Appl
5	2.0	14	9	US-09-880-748-2498	Sequence 2498, Ap
5	2.0	15	9	US-10-012-542-531	Sequence 531, Appl
5	2.0	15	9	US-10-157-669-36	Sequence 36, Appl
5	2.0	16	9	US-09-992-331-34	Sequence 34, Appl
5	2.0	16	10	US-09-879-957-167	Sequence 167, Appl
5	2.0	18	9	US-09-820-096B-5	Sequence 5, Appl
5	2.0	18	9	US-09-880-748-3031	Sequence 3031, Ap
5	2.0	18	10	US-09-820-296-5	Sequence 5, Appl
5	2.0	20	9	US-09-974-879-285	Sequence 285, Appl
5	2.0	20	9	US-10-044-995-17	Sequence 17, Appl
5	2.0	20	10	US-09-864-761-33981	Sequence 33981, A
5	2.0	20	10	US-09-864-761-43303	Sequence 43303, A
5	2.0	20	10	US-09-941-611-17	Sequence 17, Appl
5	2.0	20	12	US-10-041-030-26	Sequence 26, Appl
5	2.0	21	9	US-09-983-802-525	Sequence 525, Appl
5	2.0	21	10	US-09-027-956-6	Sequence 6, Appl
5	2.0	21	10	US-09-864-761-33828	Sequence 33828, A
5	2.0	21	12	US-10-001-843-160	Sequence 160, Appl
5	2.0	22	10	US-09-864-761-43909	Sequence 43909, A
5	2.0	23	10	US-10-097-065-286	Sequence 286, Appl
5	2.0	23	10	US-09-864-761-40975	Sequence 40975, A
5	2.0	23	10	US-09-853-253-5	Sequence 5, Appl
5	2.0	23	10	US-09-853-253-6	Sequence 6, Appl
5	2.0	23	10	US-09-759-387A-5	Sequence 5, Appl
5	2.0	24	9	US-09-991-548-3	Sequence 3, Appl
5	2.0	24	9	US-09-764-872-326	Sequence 326, Appl
5	2.0	24	10	US-09-864-761-44591	Sequence 44591, A
5	2.0	24	10	US-09-853-253-4	Sequence 4, Appl
5	2.0	25	9	US-09-986-480-226	Sequence 226, Appl
5	2.0	25	10	US-09-802-853-2	Sequence 2, Appl
5	2.0	25	10	US-09-803-165-30	Sequence 30, Appl
5	2.0	25	12	US-10-066-151-32	Sequence 32, Appl
5	2.0	27	9	US-09-991-548-36	Sequence 36, Appl
5	2.0	27	9	US-09-925-299-1289	Sequence 1289, Ap
5	2.0	27	10	US-09-864-761-34266	Sequence 34266, A
5	2.0	27	10	US-09-864-761-46105	Sequence 46105, A
5	2.0	27	10	US-09-864-761-49508	Sequence 49508, A
5	2.0	27	10	US-09-925-299-1289	Sequence 1289, Ap
5	2.0	28	9	US-09-785-059-1	Sequence 1, Appl
5	2.0	28	9	US-09-880-132-7	Sequence 7, Appl
5	2.0	28	9	US-10-079-075-1	Sequence 1, Appl
5	2.0	28	9	US-09-974-879-342	Sequence 342, Appl
5	2.0	28	9	US-09-785-058-1	Sequence 1, Appl
5	2.0	28	10	US-10-068-564-23	Sequence 23, Appl
5	2.0	28	10	US-09-739-907-172	Sequence 172, Appl
5	2.0	28	10	US-09-864-761-37297	Sequence 37297, A
5	2.0	28	10	US-09-864-761-48721	Sequence 48721, A
5	2.0	28	10	US-09-989-903-23	Sequence 23, Appl
5	2.0	28	10	US-09-880-149-7	Sequence 7, Appl
5	2.0	29	10	US-09-864-761-40217	Sequence 40217, A
5	2.0	29	10	US-09-957-607-11	Sequence 11, Appl
5	2.0	30	9	US-09-966-782A-25	Sequence 25, Appl
5	2.0	30	10	US-09-864-761-36009	Sequence 36009, A
5	2.0	31	9	US-09-785-059-2	Sequence 2, Appl
5	2.0	31	9	US-10-079-075-2	Sequence 2, Appl
5	2.0	31	9	US-10-079-075-3	Sequence 3, Appl

93 Sequence 2, Appli
 94 Sequence 3, Appli
 95 Sequence 39237, A
 96 Sequence 47417, A
 97 Sequence 48437, A
 98 Sequence 42602, A
 99 Sequence 38400, A
 100 Sequence 43639, A
 101 Sequence 45536, A
 102 Sequence 12, Appl
 103 Sequence 296, App
 104 Sequence 124, App
 105 Sequence 34574, A
 106 Sequence 42152, A
 107 Sequence 256, App
 108 Sequence 1039, App
 109 Sequence 38047, A
 110 Sequence 42216, A
 111 Sequence 47226, A
 112 Sequence 33333, A
 113 Sequence 35959, A
 114 Sequence 38517, A
 115 Sequence 6, Appl
 116 Sequence 216, App
 117 Sequence 723, App
 118 Sequence 34631, A
 119 Sequence 37614, A
 120 Sequence 39261, A
 121 Sequence 2, Appl
 122 Sequence 415, App
 123 Sequence 33530, A
 124 Sequence 171, App
 125 Sequence 1455, App
 126 Sequence 37024, A
 127 Sequence 46179, A
 128 Sequence 163, App
 129 Sequence 42062, A
 130 Sequence 123, App
 131 Sequence 517, App
 132 Sequence 517, App
 133 Sequence 36082, A
 134 Sequence 38, Appl
 135 Sequence 45161, A

AL: COMMENTS

RESULT 1
 US-09-809-638 44
 ; Sequence 44, Application US/09809638
 ; Publication No. US20030059895A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Faris
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Steve Chappell Mitchell
 ; APPLICANT: Daniel E.H. Afar
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Aya Jakobovits
 ; TITLE OF INVENTION: 125P5C8: A TISSUE SPECIFIC PROTEIN
 ; FILE REFERENCE: 129.350S01
 ; CURRENT APPLICATION NUMBER: US/09/809,638
 ; CURRENT FILING DATE: 2001-03-14
 ; NUMBER OF SEQ ID NOS: 746
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 44
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-809-638-44

Query Match 2.4%; Score 6; DB 9; Length 9;

Best Local Similarity 100.0%; Pred. No. 2.2e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AELSDS 219
 DB 2 AELSDS 7

RESULT 2

US-09-809-638-99
 ; Sequence 99, Application US/09809638
 ; Publication No. US20030059895A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Faris
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Steve Chappell Mitchell
 ; APPLICANT: Daniel E.H. Afar
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Aya Jakobovits
 ; TITLE OF INVENTION: 125P5C8: A TISSUE SPECIFIC PROTEIN
 ; FILE REFERENCE: 129.350S01
 ; CURRENT APPLICATION NUMBER: US/09/809,638
 ; CURRENT FILING DATE: 2001-03-14
 ; NUMBER OF SEQ ID NOS: 746
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 99
 ; LENGTH: 10
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-809-638-99

Query Match 2.4%; Score 6; DB 9; Length 10;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AELSDS 219
 DB 2 AELSDS 7

RESULT 3

US-09-929-924-12
 ; Sequence 12, Application US/09929924
 ; Patent No. US20020147151A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Center, David M.
 ; APPLICANT: Cruickshank, William W.
 ; APPLICANT: Kornfeld, Hardy
 ; TITLE OF INVENTION: IL-16 ANTAGONISTS
 ; FILE REFERENCE: Research Corporation Tech., Inc.
 ; CURRENT APPLICATION NUMBER: US/09/929,924
 ; CURRENT FILING DATE: 2001-08-15
 ; PRIOR APPLICATION NUMBER: 09/368,632
 ; PRIOR FILING DATE: 1999-08-05
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 16
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; INFORMATION: Description of Artificial Sequence: IL-16 antagonist peptide
 US-09-929-924-12

Query Match 2.4%; Score 6; DB 10; Length 16;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSQ 221
 DB 7 LSDSQ 12

```

RESULT 4
US-09-804-65-27
; Sequence 27, Application US/99803155
; Patent No. US20020052036A1
; GENERAL INFORMATION:
; APPLICANT: Sobek, Harold
; APPLICANT: Frey, Bruno
; APPLICANT: Antarkian, Sarafed
; APPLICANT: Boehlke, Kristina
; APPLICANT: Pisani, Francesca Maria
; APPLICANT: Kossli, Mose
; TITLE OF INVENTION: Mutant B-type DNA Polymerases Exhibiting Improved Performance
; FILE REFERENCE: 5328
; CURRENT APPLICATION NUMBER: US/09/803.165
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: EP/00105155.6
; PRIOR FILING DATE: 2000-03-11
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 25
; TYPE: PRT
; ORGANISM: S. solfatarius
US-09-803-155-27

```

```

Query Match      2.4%  Score 6;  DB 10;  Length 25;
Best Local Similarity 100.0%;  Pred. No. 78;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY  25 SALING 30
    |||||
DB   2 SALING 7

```

```

RESULT 5
US-09-864-761-49063
; Sequence 49063, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENE-DRIVEN SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Reomica X-1
; CURRENT APPLICATION NUMBER: US/09/864.761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: CH 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,667
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anconax Sequence Listing Engine vers. 1.1
; SEQ ID NO 49063
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007597.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 0.6
; OTHER INFORMATION: EST HUMAN HIT: AW086043.1, EVALU 1.00e-04
; OTHER INFORMATION: SWISSPROT HIT: P53632, EVALU 1.40e-01
US-09-864-761-49063

```

```

Query Match      2.4%  Score 6;  DB 10;  Length 29;
Best Local Similarity 100.0%;  Pred. No. 89;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY  178 VNRIES 183
    |||||
DB   14 VNRIES 19

```

```

RESULT 6
US-09-142-755-9
; Sequence 9, Application US/09142755
; Patent No. US20020076725A1
; GENERAL INFORMATION:
; APPLICANT: TOYOSAKI-MAEDA, Tomoko
; APPLICANT: SUZUKI, Ryuji
; APPLICANT: TSURUTA, Yuji
; APPLICANT: TAKEMOTO, Hiroshi
; TITLE OF INVENTION: HUMAN T CELL CLONES SPECIFIC FOR RHEUMATOID ARTHRITIS
; FILE REFERENCE: 74129/452
; CURRENT APPLICATION NUMBER: US/09/142.755
; CURRENT FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: PCT/JPS7/00774
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: JP 8/56022
; PRIOR FILING DATE: 1996-03-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-142-755-9

```

```

Query Match      2.4%  Score 6;  DB 10;  Length 32;
Best Local Similarity 100.0%;  Pred. No. 97;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY  32 TRPYE 37
    |||||
DB   15 TRPYE 20

```

```

RESULT 7
US-09-843-676-178
; Sequence 178, Application US/09843676
; Patent No. US20020164788A1
; GENERAL INFORMATION:

```

APPLICANT: Cech, Thomas R.
 Lingner, Joachim
 Nakamura, Toru
 Chapman, Karen B.
 Morin, Gregg B.
 Harley, Calvin
 Andrews, William H.
 TITLE OF INVENTION: No. US20020164786A1el Telomerase
 NUMBER OF SEQUENCES: 225
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/843,676
 FILING DATE: 26-APR-2001
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/854,350
 FILING DATE: 09-MAY-1997
 APPLICATION NUMBER: US/08/846,017
 FILING DATE: 25-APR-1997
 APPLICATION NUMBER: US/08/844,419
 FILING DATE: 18-APR-1997
 APPLICATION NUMBER: US/08/724,643
 FILING DATE: 01-OCT-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-0029300S
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 178:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 35 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 178:
 US-09-843-676-178

Query Match 2.4% Score 6; DB 9; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 243 GAUGIA 248
 Db 10 GAUGIA 15

RESULT 8
 US-09-438-486-178
 Sequence 178, Application US/09/438486
 Publication No. US20030009019A1
 GENERAL INFORMATION:
 APPLICANT: Cech, Thomas R.
 APPLICANT: Lingner, Joachim
 APPLICANT: Nakamura, Toru
 APPLICANT: Chapman, Karen B.
 APPLICANT: Morin, Gregg B.
 APPLICANT: Harley, Calvin
 APPLICANT: Andrews, William H.
 TITLE OF INVENTION: No. US20030009019A1el Telomerase

NUMBER OF SEQUENCES: 223
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/438,486
 FILING DATE: 12-NOV-1999
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/851,843
 FILING DATE: 06-MAY-1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/844,419
 FILING DATE: 18-APR-1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/724,643
 FILING DATE: 01-OCT-1996
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-0029310S
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 178:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 35 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-438-486-178

Query Match 2.4% Score 6; DB 9; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 243 GAUGIA 248
 Db 10 GAUGIA 15

RESULT 9
 US-10-053-758-178
 Sequence 178, Application US/10053758
 Publication No. US20030032075A1
 GENERAL INFORMATION:
 APPLICANT: Cech, Thomas R.
 APPLICANT: Lingner, Joachim
 APPLICANT: Nakamura, Toru
 APPLICANT: Chapman, Karen B.
 APPLICANT: Morin, Gregg B.
 APPLICANT: Harley, Calvin
 APPLICANT: Andrews, William H.
 TITLE OF INVENTION: No. US20030032075A1el Telomerase
 NUMBER OF SEQUENCES: 225
 CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/053,758
; FILING DATE: 18-Jan-2002
; CLASSIFICATION: 536
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 0:5389 002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-10-053-758-178

Query Match 2.4%; Score 6; DA 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 GAAGLA 248
Db 10 GAAGLA 15

RESULT 10
US-10-054-295-178
; Sequence 178, Application US/10054295
; Publication No. US20030044953A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
;
; TITLE OF INVENTION: No. US20030044953A1el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,611
; FILING DATE: 18-Jan-2002

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,295
; FILING DATE: 18-Jan-2002
; CLASSIFICATION: 536
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/854,050
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 0:5389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-10-054-295-178

Query Match 2.4%; Score 6; DA 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 GAAGLA 248
Db 10 GAAGLA 15

RESULT 11
US-10-054-611-178
; Sequence 178, Application US/10054611
; Publication No. US20030059787A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
;
; TITLE OF INVENTION: No. US20030059787A1el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,611
; FILING DATE: 18-Jan-2002

```

CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/854,050
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 08/846,017
 FILING DATE: 25-APR-1997
 APPLICATION NUMBER: US 08/844,419
 FILING DATE: 18-APR-1997
 APPLICATION NUMBER: US 08/724,643
 FILING DATE: 01-OCT-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Applc. Randolph I.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-00293005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 178:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 35 amino acids
 TYPE: amino acid
 STRANDEDNESS: linear
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 178:
 US-10-054-611-178

Query Match 2.4%; Score 6; DB 9; Length 35;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 GAGGLA 248
 |||||
 DB 10 GAGGLA 15

RESULT 12
 US-09-864 761-46227
 ; Sequence 46227, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Weiseng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aeonica X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,455
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 46227
 ; LENGTH: 39
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC008655.3
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 40
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 1.3
 ; OTHER INFORMATION: EST HUMAN HIT: BF513122.1, EVALU 6.00e-16
 ; OTHER INFORMATION: SWISSPROT HIT: O96009, EVALU 5.00e-17
 US-09-864-761-46227

Query Match 2.4%; Score 6; DB 10; Length 39;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 202 PKASSS 207
 |||||
 DB 7 PKASSS 12

RESULT 13
 US-09-989-919-101
 ; Sequence 101, Application US/09989919
 ; Patent No. US2002016344A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Recipon, Hervé
 ; APPLICANT: Pluta, Jason
 ; APPLICANT: Ghosh, Malavika
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Chenghua
 ; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and
 ; FILE REFERENCE: DEX-0289
 ; CURRENT APPLICATION NUMBER: US/09/989,919
 ; CURRENT FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: 60/252,505
 ; PRIOR FILING DATE: 2000-11-22
 ; NUMBER OF SEQ ID NOS: 124
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 10:
 ; LENGTH: 61
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 US-09-989-919-101

Query Match 2.4%; Score 6; DB 9; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 KASSSP 208
 |||||
 DB 3 KASSSP 8

RESULT 14
 US-09-864-761-44746
 ; Sequence 44746, Application US/09864761


```

1 Patent No. US20020048763A1
2 GENERAL INFORMATION:
3 APPLICANT: Penn, Sharron G.
4 APPLICANT: Rank, David R.
5 APPLICANT: Hanzel, David K.
6 APPLICANT: Chen, Wensheng
7 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
8 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
9 FILE REFERENCE: Acomica-X-1
10 CURRENT APPLICATION NUMBER: US/09/864,761
11 CURRENT FILING DATE: 2001-05-23
12 PRIOR APPLICATION NUMBER: US 60/190,312
13 PRIOR FILING DATE: 2000-02-04
14 PRIOR APPLICATION NUMBER: US 60/227,456
15 PRIOR FILING DATE: 2000-05-26
16 PRIOR APPLICATION NUMBER: US 09/632,366
17 PRIOR FILING DATE: 2000-08-03
18 PRIOR APPLICATION NUMBER: GB 24263.6
19 PRIOR FILING DATE: 2000-10-04
20 PRIOR APPLICATION NUMBER: US 60/236,359
21 PRIOR FILING DATE: 2000-09-27
22 PRIOR APPLICATION NUMBER: PCT/US01/00666
23 PRIOR FILING DATE: 2001-01-30
24 PRIOR APPLICATION NUMBER: PCT/US01/00667
25 PRIOR FILING DATE: 2001-01-30
26 PRIOR APPLICATION NUMBER: PCT/US01/00668
27 PRIOR FILING DATE: 2001-01-30
28 PRIOR APPLICATION NUMBER: PCT/US01/00669
29 PRIOR FILING DATE: 2001-01-30
30 PRIOR APPLICATION NUMBER: PCT/US01/00665
31 PRIOR FILING DATE: 2001-01-30
32 PRIOR APPLICATION NUMBER: PCT/US01/00668
33 PRIOR FILING DATE: 2001-01-30
34 PRIOR APPLICATION NUMBER: PCT/US01/00663
35 PRIOR FILING DATE: 2001-01-30
36 PRIOR APPLICATION NUMBER: PCT/US01/00662
37 PRIOR FILING DATE: 2001-01-30
38 PRIOR APPLICATION NUMBER: PCT/US01/00661
39 PRIOR FILING DATE: 2001-01-30
40 PRIOR APPLICATION NUMBER: PCT/US01/00670
41 PRIOR FILING DATE: 2001-01-30
42 PRIOR APPLICATION NUMBER: US 60/234,687
43 PRIOR FILING DATE: 2000-09-21
44 PRIOR APPLICATION NUMBER: US 09/608,408
45 PRIOR FILING DATE: 2000-06-30
46 PRIOR APPLICATION NUMBER: US 09/774,203
47 PRIOR FILING DATE: 2001-01-29
48 NUMBER OF SEQ ID NOS: 49117
49 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
50 SEQ ID NO 44746
51 LENGTH: 61
52 TYPE: PRT
53 ORGANISM: Homo sapiens
54 FEATURE:
55 OTHER INFORMATION: MAP TO AC007923.2
56 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
57 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7
58 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
59 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
60 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
61 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
62 OTHER INFORMATION: EST_HUMAN HIT: BEIG8410.1, EVALUATE 3.00e-14
63 OTHER INFORMATION: SWISSPROT HIT: P03355, EVALUATE 7.00e-09
64 US-09 864-761-44746

```

```

Query Match 2.48; Score 6; DB 10; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.7e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 ILA1SL 8
   11111
DB 23 ILA1SL 28

```

```

RESULT 15
US-09-815-242-13075
Sequence 13075, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA-011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13075
LENGTH: 61
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-13075

```

```

Query Match 2.48; Score 6; DB 10; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.7e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 193 KAAALP 198
   111111
DB 23 KAAALP 28

```

```

RESULT 16
US-09-864-761-42526
Sequence 42526, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Acomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359

```

; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO: 42526
 ; LENGTH: 63
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC024460.1
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 1.9
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 2
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 2.7
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 4.1
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 2.6
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 2.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 4.5
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1.8
 ; OTHER INFORMATION: SWISSPROT HIT: Q08379, EVALUE 2.00e-03
 ; OTHER INFORMATION: EST_HUMAN HIT: BE087405.1, EVALUE 2.00e-07
 US-09-864-761-42526

Query Match 2.4%; Score 6; DB 10; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 PEPPAV 213
 DB 3 PEPPAV 8

RESULT 17
 US-09-864-761-42806
 ; Sequence 42806, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon C.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Ranzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Acornica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-09-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO: 42806
 ; LENGTH: 65
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO 298752.9
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL - 1.1
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 2.4
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 0.99
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 1.6
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 1.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 3
 ; OTHER INFORMATION: EST_HUMAN HIT: AA380168.1, EVALUE 6.00e-27
 ; OTHER INFORMATION: SWISSPROT HIT: Q01538, EVALUE 3.00e-10
 US-09-864-761-42806

Query Match 2.4%; Score 6; DB 10; Length 65;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AELSDS 219
 DB 38 AELSDS 43

RESULT 18
 US-10-092-154-629
 ; Sequence 629, Application US/10092154
 ; Publication No. US20030054375A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC009C1
 ; CURRENT APPLICATION NUMBER: US/10/092,154
 ; CURRENT FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 2003
 ; Prior Application removed - See File Wrapper or Palm

```
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 629
? LENGTH: 69
? TYPE: PKT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (110)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: misc_feature
? LOCATION: (113)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: misc_feature
? LOCATION: (40)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: misc_feature
? LOCATION: (62)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10 092-154-629
Query Match 2.4%; Score 6; DB 9; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 GAGGCG 133
Db 43 GAGGCG 48

RESULT 19
US-09-764-847-629
? Sequence 629, Application US/03764847
? Patent No. US20020132767A1
? GENERAL INFORMATION:
? APPLICANT: Roser et al.
? TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
? FILE REFERENCE: PC009
? CURRENT APPLICATION NUMBER: US/05/764,847
? CURRENT FILING DATE: 2001-01-17
? Prior application data removed - consult PALM or file wrapper
? NUMBER OF SEQ ID NOS: 2003
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 629
? LENGTH: 69
? TYPE: PKT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (13)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (113)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (40)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (62)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764 847-629
Query Match 2.4%; Score 6; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 GAGGCG 133
Db 43 GAGGCG 48

RESULT 20
US-09-799-118-8
? Sequence 8, Application US/03799118
```

```
? Patent No. US20020090708A1
? GENERAL INFORMATION:
? APPLICANT: Lee, Tae Ho
? Wilniewski, Hans Georg
? Wilcek, Jan
? TITLE OF INVENTION: Cytokine-Induced Protein, ISG-6, DNA Coding
? Therefor and Uses Thereof
? NUMBER OF SEQUENCES: 11
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Browdy and Neimark
? STREET: 419 Seventh Street, N.W., Suite 300
? CITY: Washington
? STATE: D.C.
? ZIP: 20004
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/799,118
? FILING DATE: 06-Mar-2001
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/242,097
? FILING DATE: 13-MAY-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: BROWDY, Roger L.
? REGISTRATION NUMBER: 25,618
? REFERENCE/DOCKET NUMBER: LEF26/VILCEK-18
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 212-628-5197
? TELEFAX: 212-737-3528
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 5 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-799-118-8
Query Match 2.0%; Score 5; DB 10; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SLPEG 11
Db 1 SLPEG 5

RESULT 21
US-09-859-214-41
? Sequence 41, Application US/09859214
? Patent No. US20020103111A1
? GENERAL INFORMATION:
? APPLICANT: Schwender, Charles F.
? Shroff, Hitesh N.
? TITLE OF INVENTION: INHIBITORS OF MADCAM-1-MEDIATED
? INTERACTIONS AND METHODS OF USE THEREFOR
? NUMBER OF SEQUENCES: 89
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
? STREET: Two Millitia Drive
? CITY: Lexington
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02421
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/99/859,214
 ; FILING DATE: 16-May-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 69/109,879
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: US 68/562,740
 ; FILING DATE: 04-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brook, David E.
 ; REGISTRATION NUMBER: 22,592
 ; REFERENCE/BOOK NUMBER: LKS95-12A2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781) 861-6240
 ; TELEFAX: (781) 861-9540
 ; INFORMATION FOR SEQ ID NO: 41:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: Modified-site
 ; LOCATION: 1
 ; OTHER INFORMATION: /label= modified aa
 ; /note= "Pda - Leucine"
 ; FEATURE:
 ; NAME/KEY: Modified-site
 ; LOCATION: 5
 ; OTHER INFORMATION: /label= modified aa
 ; /note= "Leucine - NH2"
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 41:
 US-09-859 214-41

Query Match 2.0%; Score 5; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LATSLSR
 ;
 DB 1 LATSLS

RESULT 22
 ; Sequence 213, Application US/09989789
 ; Patent No. US20020063379A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LIU, Qiang
 ; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
 ; FILE REFERENCE: 8325-0011.20 / S11-US2
 ; CURRENT APPLICATION NUMBER: US/09/989,789
 ; CURRENT FILING DATE: 2002-03-25
 ; NUMBER OF SEQ ID NOS: 4085
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 213
 ; LENGTH: 7
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
 US-09 989-789-213

Query Match 2.0%; Score 5; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 SSSLQ 102
 ;
 DB 11111

DB 2 SSSLQ 6

RESULT 23
 ; Sequence 214, Application US/09989789
 ; Patent No. US20020063379A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LIU, Qiang
 ; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
 ; FILE REFERENCE: 8325-0011.20 / S11-US2
 ; CURRENT APPLICATION NUMBER: US/09/989,789
 ; CURRENT FILING DATE: 2002-03-25
 ; NUMBER OF SEQ ID NOS: 4085
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 214
 ; LENGTH: 7
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
 US-09-989-789-214

Query Match 2.0%; Score 5; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 SSSLQ 102
 ;
 DB 11111

RESULT 24
 ; Sequence 222, Application US/09989789
 ; Patent No. US20020063379A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LIU, Qiang
 ; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
 ; FILE REFERENCE: 8325-0011.20 / S11-US2
 ; CURRENT APPLICATION NUMBER: US/09/989,789
 ; CURRENT FILING DATE: 2002-03-25
 ; NUMBER OF SEQ ID NOS: 4085
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 222
 ; LENGTH: 7
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
 US-09-989-789-222

Query Match 2.0%; Score 5; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 SSSLQ 102
 ;
 DB 11111

RESULT 25
 ; Sequence 16, Application US/09945917
 ; Publication No. US2003004238:A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bogaert, Thierry
 ; APPLICANT: Vandekerckhove, Joel
 ; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
 ; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR

; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
 ; FILE REFERENCE: P/14-1
 ; CURRENT APPLICATION NUMBER: US/945,917
 ; CURRENT FILING DATE: 1998-09-21
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 16
 ; LENGTH: 8
 ; TYPE: PR1
 ; ORGANISM: Caenorhabditis elegans
 US-09-945-917-16

Query Match 2.0%; Score 5; DB 9; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 52 PAVPP 56
 Db 1 PAVPP 5

RESULT 26
 US-10-125-635A-93
 ; Sequence 93, Application US/10125635A
 ; Publication No. US20030039635A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Galiger, Alexander
 ; APPLICANT: Smithgall, Molly D.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Sutherland, R. Alec
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 ; FILE REFERENCE: 210-21,46507
 ; CURRENT APPLICATION NUMBER: US/10125,635A
 ; CURRENT FILING DATE: 2002-07-19
 ; NUMBER OF SEQ ID NOS: 461
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 93
 ; LENGTH: 9
 ; TYPE: PR1
 ; ORGANISM: Homo sapiens
 US-10-125-635A-93

Query Match 2.0%; Score 5; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 227 GGPAP 231
 Db 1 GGPAP 5

RESULT 27
 US-09-945-917-33
 ; Sequence 33, Application US/99945917
 ; Publication No. US2003004238A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Boogaert, Thierry
 ; APPLICANT: Vandekerckhove, Joel
 ; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
 ; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
 ; TITLE OF INVENTION: PHARMACOLOGICAL COMPOSITIONS CONTAINING THEM AND THEIR
 ; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
 ; FILE REFERENCE: P/14-1
 ; CURRENT APPLICATION NUMBER: US/09/945,917
 ; CURRENT FILING DATE: 1998-09-21
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 33
 ; LENGTH: 9
 ; TYPE: PR1

; ORGANISM: Caenorhabditis elegans
 US-09-945-917-33

Query Match 2.0%; Score 5; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 52 PAVPP 56
 Db 4 PAVPP 8

RESULT 28
 US-09-809-638-645
 ; Sequence 645, Application US/09809638
 ; Publication No. US20030059895A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Faris
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Steve Chappell Mitchell
 ; APPLICANT: Daniel E.H. Afar
 ; APPLICANT: Arthur H. Raitano
 ; APPLICANT: Aya Jakobovits
 ; TITLE OF INVENTION: 125P5C8: A TISSUE SPECIFIC PROTEIN
 ; TITLE OF INVENTION: HIGHLY EXPRESSED IN VARIOUS CANCERS
 ; FILE REFERENCE: 129.35CS01
 ; CURRENT APPLICATION NUMBER: US/09/809,638
 ; CURRENT FILING DATE: 2001-03-14
 ; NUMBER OF SEQ ID NOS: 746
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 645
 ; LENGTH: 9
 ; TYPE: PR1
 ; ORGANISM: Homo sapiens
 US-09-809-638-645

Query Match 2.0%; Score 5; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 215 ELSDS 219
 Db 1 ELSDS 5

RESULT 29
 US-09-996-288-185
 ; Sequence 185, Application US/09996288
 ; Patent No. US2002017126A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, James
 ; APPLICANT: Leslie, Johnson
 ; APPLICANT: Scott, Koenig
 ; TITLE OF INVENTION: Methods of Administering/Dosing Anti-RSV Antibodies for Prophylaxis
 ; FILE REFERENCE: 10271-047-999
 ; CURRENT APPLICATION NUMBER: US/09/996,288
 ; CURRENT FILING DATE: 2001-11-28
 ; NUMBER OF SEQ ID NOS: 259
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 185
 ; LENGTH: 10
 ; TYPE: PR1
 ; ORGANISM: Homo sapiens
 US-09-996-288-185

Query Match 2.0%; Score 5; DB 9; Length 10;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 203 KASSS 207
 Db 1 KASSS 5

TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
ANTIBODIES TO HEPATITIS C VIRUS

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/044,995
FILING DATE: 15-Jan-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/391,671
FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCI-1992

APPLICATION NUMBER: WC PCI/EP91/02409
FILING DATE: 13-DEC-1991

APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990

ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5

TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-044-995-22

Query Match 2.0%; Score 5; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 AQUALP 198
Db 3 AQUALP 7

RESULT 33
US-09-941-611-22
Sequence 22, Application US/09941611
Patent No. US2002006640A1
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
POLLET, DIRK
MAERTENS, GEERT
VAN HEUVERSWJN, HUGO

TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
ANTIBODIES TO HEPATITIS C VIRUS

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/044,995
FILING DATE: 15-Jan-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/391,671
FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCI-1992

APPLICATION NUMBER: WC PCI/EP91/02409
FILING DATE: 13-DEC-1991

APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990

ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5

TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-044-995-22

Query Match 2.0%; Score 5; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 AQUALP 198
Db 3 AQUALP 7

RESULT 33
US-09-941-611-22
Sequence 22, Application US/09941611
Patent No. US2002006640A1
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
POLLET, DIRK
MAERTENS, GEERT
VAN HEUVERSWJN, HUGO

TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
ANTIBODIES TO HEPATITIS C VIRUS

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/044,995
FILING DATE: 15-Jan-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/391,671
FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCI-1992

APPLICATION NUMBER: WC PCI/EP91/02409
FILING DATE: 13-DEC-1991

APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990

ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5

TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-044-995-22

Query Match 2.0%; Score 5; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 AQUALP 198
Db 3 AQUALP 7

RESULT 33
US-09-941-611-22
Sequence 22, Application US/09941611
Patent No. US2002006640A1
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
POLLET, DIRK
MAERTENS, GEERT
VAN HEUVERSWJN, HUGO

TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
ANTIBODIES TO HEPATITIS C VIRUS

NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/044,995
FILING DATE: 15-Jan-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/391,671
FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCI-1992

APPLICATION NUMBER: WC PCI/EP91/02409
FILING DATE: 13-DEC-1991

APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990

ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5

TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-044-995-22

Query Match 2.0%; Score 5; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 194 AQUALP 198
Db 3 AQUALP 7

RESULT 33
US-09-941-611-22
Sequence 22, Application US/09941611
Patent No. US2002006640A1
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
POLLET, DIRK
MAERTENS, GEERT
VAN HEUVERSWJN, HUGO

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/941,611
 FILING DATE: 30-Aug-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 98/391,671
 FILING DATE: 1995-02-21
 APPLICATION NUMBER: WO PCT/EP91/02409
 FILING DATE: 13-DEC-1991
 APPLICATION NUMBER: EP 90124241.2
 FILING DATE: 14-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: SADOFF, B.J.
 REGISTRATION NUMBER: 36,663
 REFERENCE/DOCKET NUMBER: 1487-5
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 7038164000
 TELEFAX: 7038164100
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 22:
 US-09-941-611:22

Query Match 2.0%; Score 5; DB 10; Length 11;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 AQAAP 158
 DB 3 AQAAP 7

RESULT 34
 US-10-125-635A 310
 : Sequence 310, Application US/0025635A
 : Publication No. US20030009635A1
 : GENERAL INFORMATION:
 : APPLICANT: Galger, Alexander
 : APPLICANT: Smithall, Molly D.
 : APPLICANT: Carter, Darick
 : APPLICANT: Cheever, Martin A.
 : APPLICANT: McNeill, Patricia D.
 : APPLICANT: Sutherland, R. Alec
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
 : TITLE OF INVENTION: SPECIFIC IMMUNOTHERAPY
 : FILE REFERENCE: 210121.465C7
 : CURRENT APPLICATION NUMBER: US/10/125,635A
 : CURRENT FILING DATE: 2002-07-19
 : NUMBER OF SEQ ID NOS: 461
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 310
 : LENGTH: 12
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-10-125-635A-310

Query Match 2.0%; Score 5; DB 9; Length 12;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 GCIAP 231
 DB 8 GGAP 12

RESULT 35
 US-09-945-917-8
 : Sequence 8, Application US/09945917
 : Publication No. US20030042381A1
 : GENERAL INFORMATION:
 : APPLICANT: Bogaert, Thierry
 : APPLICANT: Vandekerckhove, Joel
 : TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
 : TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
 : TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
 : FILE REFERENCE: P/14-1
 : CURRENT APPLICATION NUMBER: US/09/945,917
 : CURRENT FILING DATE: 1998-09-21
 : NUMBER OF SEQ ID NOS: 58
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 8
 : LENGTH: 13
 : TYPE: PRT
 : ORGANISM: Caenorhabditis elegans
 US-09-945-917-8

Query Match 2.0%; Score 5; DB 9; Length 13;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 PAVPP 56
 DB 6 PAVPP 10

RESULT 36
 US-09-880-748-2498
 : Sequence 2498, Application US/09880748
 : Publication No. US20030059937A1
 : GENERAL INFORMATION:
 : APPLICANT: Ruben et al.
 : TITLE OF INVENTION: Antibodies that Immunoselectively Bind Blys
 : FILE REFERENCE: PP523
 : CURRENT APPLICATION NUMBER: US/09/880,748
 : CURRENT FILING DATE: 2001-06-15
 : PRIOR APPLICATION NUMBER: 60/212,210
 : PRIOR FILING DATE: 2000-06-15
 : PRIOR APPLICATION NUMBER: 60/240,816
 : PRIOR FILING DATE: 2000-10-17
 : PRIOR APPLICATION NUMBER: 60/276,248
 : PRIOR FILING DATE: 2001-03-16
 : PRIOR APPLICATION NUMBER: 60/277,379
 : PRIOR FILING DATE: 2001-03-21
 : PRIOR APPLICATION NUMBER: 60/293,499
 : PRIOR FILING DATE: 2001-05-25
 : NUMBER OF SEQ ID NOS: 3239
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 2498
 : LENGTH: 14
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-880-748-2498

Query Match 2.0%; Score 5; DB 9; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 AGEFS 125
 DB 10 AGEFS 14

RESULT 37
 US-10-012-542-531
 : Sequence 531, Application US/10012542

```

US-09-992-331-34
? Sequence 34, Application US/99992331
? Publication No. US20030022186A1
? GENERAL INFORMATION:
? APPLICANT: FEDER, JOHN N.
? APPLICANT: MINTIER, GABE
? APPLICANT: RAMANATHAN, CHANDRA S.
? APPLICANT: HAWKEN, DONALD R.
? TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGP8MY18,
? TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
? TITLE OF INVENTION: CELLS
? FILE REFERENCE: D0048NP
? CURRENT APPLICATION NUMBER: US/09/992,331
? CURRENT FILING DATE: 2001-11-14
? PRIOR APPLICATION NUMBER: 60/308,540
? PRIOR FILING DATE: 2001-07-27
? PRIOR APPLICATION NUMBER: 60/261,782
? PRIOR FILING DATE: 2001-01-16
? PRIOR APPLICATION NUMBER: 60/248,483
? PRIOR FILING DATE: 2000-11-14
? NUMBER OF SEQ ID NOS: 45
? SOFTWARE: PatentIn Ver. 2.1.
? SEQ ID NO 34
? LENGTH: 16
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence:Synthesized
US-09-992-331-34

Query Match          2.0%; Score 5; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      80 VSLIG 84
      |||||
DB      2 VSLIG 6

RESULT 4C
US-09-879-957-167
? Sequence 167, Application US/09879957
? Patent No. US20020034755A1
? GENERAL INFORMATION:
? APPLICANT: SPARKS, Andrew H.
? APPLICANT: HOFFMAN, No. US20020034755A1h
? APPLICANT: KAY, Brian K.
? APPLICANT: FOWLKES, Dana M.
? APPLICANT: MCCONNELL, Stephen J.
? TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
? TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
? USING SAME:
? NUMBER OF SEQUENCES: 227
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Pennie & Edmonds LLP
? STREET: 1155 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10036-2711
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/879,957
? FILING DATE: 13-Jun-2001
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/630,915
? FILING DATE: 03-APR-1996

```



```

: ATTORNEY/AGENT INFORMATION:
: NAME: WISTOCK, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 1101-174
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 750-9090
: TELEFAX: (212) 865-8864/5741
: TELEX: 66141 FENNIE
: INFORMATION FOR SEQ ID NO: 167:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 16 amino acids
: TYPE: amino acid
: STRANDEDNESS: <unknown>
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
: SEQUENCE DESCRIPTION: SEQ ID NO: 167:
US 09-879-957 167

```

```

Query Match      2.0%; Score 5; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 52 PAVPP 56
   IIII
Db 6 PAVTP 10

```

```

Search completed: April 10, 2003, 10:43:36
Job time : 13 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Ran on: April 10, 2003, 10:37:55 ; Search time 10.294; Seconds
(without alignments)
714.557 Million ce.1 updates/sec

Title: us-09-930-026-1

Perfect score: 250

Sequence: 1 MACIATSLPEGIWVTFDR.....PGASGSDSEGGAGLAFS 250

Scoring table: CLICO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 189198

Minimum DB seq length: 0

Maximum DB seq length: 70

Post-processing: listing first 135 summaries

Database : Issued_Patents_AA*

1: /cgn2_6/prodata/1/iaa/5A.COMB.pep.*

2: /cgn2_6/prodata/1/iaa/5B.COMB.pep.*

3: /cgn2_6/prodata/1/iaa/6A.COMB.pep.*

4: /cgn2_6/prodata/1/iaa/6B.COMB.pep.*

5: /cgn2_6/prodata/1/iaa/PCUS.COMB.pep.*

6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	7	2.8	62	1	US-07-676-647-5
2	7	2.8	62	1	US-08-449-329-5
3	7	2.8	62	2	US-08-445-073-5
4	7	2.8	62	2	US-08-585-258-5
5	7	2.8	62	4	US-09-211-590-5
6	7	2.8	62	5	PCP-US91-03896-5
7	6	2.4	6	2	US-08-672-610A-39
8	6	2.4	7	2	US-08-672-610A-40
9	6	2.4	8	2	US-08-672-610A-6
10	6	2.4	8	2	US-08-672-610A-51
11	6	2.4	12	3	US-08-257-783-9
12	6	2.4	12	5	PCP-US95-07157-9
13	6	2.4	16	3	US-08-893-526A-31
14	6	2.4	28	1	US-08-047-033-11
15	6	2.4	35	3	US-08-851-843A-178
16	6	2.4	35	4	US-08-974-549A-297
17	6	2.4	35	4	US-08-864-050-178
18	6	2.4	45	4	US-09-430-323-178
19	6	2.4	43	2	US-08-472-244-1
20	6	2.4	46	1	US-08-340-428B-36
21	6	2.4	46	5	PCP-US93-07306-36
22	6	2.4	50	1	US-08-294-189 16
23	6	2.4	67	3	US-09-120-365-94
24	6	2.4	67	3	US-09-120-365-95
25	6	2.4	67	3	US-09-120-365-96
26	6	2.4	67	3	US-09-120-365-97
27	6	2.4	67	3	US-09-120-365-98
28	6	2.4	6	3	US-09-120-365-99
29	6	2.4	6	4	US-09-515-039-94
30	6	2.4	6	4	US-09-515-039-95
31	6	2.4	6	4	US-09-515-039-96
32	6	2.4	6	4	US-09-515-039-97
33	6	2.4	6	4	US-09-515-039-98
34	6	2.4	6	4	US-09-515-039-99
35	6	2.4	70	4	US-09-134-001C-4974
36	5	2.0	5	2	US-08-242-097-8
37	5	2.0	5	2	US-08-672-610A-36
38	5	2.0	5	3	US-08-582-740-41
39	5	2.0	5	4	US-09-206-695-8
40	5	2.0	5	4	US-09-109-879-41
41	5	2.0	5	4	US-09-000-179-3
42	5	2.0	5	5	PCP-US96-11995-3
43	5	2.0	6	2	US-08-672-610A-37
44	5	2.0	7	1	US-08-442-063A-17
45	5	2.0	7	2	US-08-672-610A-38
46	5	2.0	7	2	US-08-097-554A-87
47	5	2.0	7	3	US-08-480-640A-87
48	5	2.0	7	3	US-08-295-802-87
49	5	2.0	7	4	US-08-488-237A-87
50	5	2.0	7	4	US-08-375-922A-87
51	5	2.0	8	3	US-08-336-553A-44
52	5	2.0	8	3	US-08-336-553A-60
53	5	2.0	8	3	US-08-336-553A-72
54	5	2.0	8	3	US-08-941-445A-27
55	5	2.0	8	4	US-08-444-818-578
56	5	2.0	8	4	US-08-444-818-579
57	5	2.0	8	4	US-08-444-818-580
58	5	2.0	8	4	US-08-439-157-44
59	5	2.0	8	4	US-08-439-157-60
60	5	2.0	8	4	US-08-439-157-72
61	5	2.0	8	4	US-09-437-895-44
62	5	2.0	8	4	US-09-437-895-60
63	5	2.0	8	4	US-09-437-895-72
64	5	2.0	9	1	US-08-164-839-45
65	5	2.0	9	1	US-08-583-759-45
66	5	2.0	9	1	US-08-102-738-11
67	5	2.0	9	2	US-08-146-028-409
68	5	2.0	9	2	US-08-146-028-410
69	5	2.0	9	2	US-08-146-028-411
70	5	2.0	9	2	US-08-146-028-412
71	5	2.0	9	2	US-08-146-028-413
72	5	2.0	9	3	US-08-159-339A-819
73	5	2.0	9	3	US-08-159-339A-1220
74	5	2.0	9	4	US-08-723-425A-409
75	5	2.0	9	4	US-08-723-425A-410
76	5	2.0	9	4	US-08-723-425A-411
77	5	2.0	9	4	US-08-723-425A-412
78	5	2.0	9	4	US-08-723-425A-413
79	5	2.0	9	4	US-09-112-206-409
80	5	2.0	9	4	US-09-112-206-410
81	5	2.0	9	4	US-09-112-206-411
82	5	2.0	9	4	US-09-112-206-412
83	5	2.0	9	4	US-09-112-206-413
84	5	2.0	10	1	US-08-164-839-66
85	5	2.0	10	1	US-08-583-799-66
86	5	2.0	10	2	US-08-572-951-33
87	5	2.0	10	2	US-08-518-967-7
88	5	2.0	10	4	US-09-537-357-1
89	5	2.0	10	6	5314872-3
90	5	2.0	10	6	5464943-13
91	5	2.0	11	1	US-08-336-343A-27
92	5	2.0	11	2	US-08-466-975A-22
93	5	2.0	11	2	US-08-391-671A-22
94	5	2.0	11	2	US-08-765-745-24
95	5	2.0	11	3	US-08-467-902A-22
96	5	2.0	11	4	US-08-652-877-27
97	5	2.0	11	4	US-08-476-515A-27
98	5	2.0	11	4	US-09-275-265-22
99	5	2.0	12	1	US-08-260-582-50
100	5	2.0	12	1	US-08-442-063A-19

Sequence 99, Appl
Sequence 94, Appl
Sequence 95, Appl
Sequence 96, Appl
Sequence 97, Appl
Sequence 98, Appl
Sequence 99, Appl
Sequence 4974, Ap
Sequence 8, Appl
Sequence 36, Appl
Sequence 41, Appl
Sequence 8, Appl
Sequence 41, Appl
Sequence 3, Appl
Sequence 37, Appl
Sequence 17, Appl
Sequence 38, Appl
Sequence 87, Appl
Sequence 87, Appl
Sequence 87, Appl
Sequence 87, Appl
Sequence 87, Appl
Sequence 44, Appl
Sequence 60, Appl
Sequence 72, Appl
Sequence 27, Appl
Sequence 578, App
Sequence 579, App
Sequence 580, App
Sequence 44, Appl
Sequence 60, Appl
Sequence 72, Appl
Sequence 44, Appl
Sequence 60, Appl
Sequence 72, Appl
Sequence 45, Appl
Sequence 45, Appl
Sequence 11, Appl
Sequence 409, App
Sequence 410, App
Sequence 411, App
Sequence 412, App
Sequence 413, App
Sequence 819, App
Sequence 1220, Ap
Sequence 409, App
Sequence 410, App
Sequence 411, App
Sequence 412, App
Sequence 413, App
Sequence 409, App
Sequence 410, App
Sequence 411, App
Sequence 412, App
Sequence 413, App
Sequence 66, Appl
Sequence 66, Appl
Sequence 33, Appl
Sequence 7, Appl
Sequence 1, Appl
Patent No. 5314872
Patent No. 5464943
Sequence 27, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 24, Appl
Sequence 22, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 50, Appl
Sequence 19, Appl

```

101 5 2.0 12 1 US-08-433-854-21
102 5 2.0 12 1 US-08-174-745A-21
103 5 2.0 12 1 US-08-195-947-21
104 5 2.0 12 2 US-08-433-885-21
105 5 2.0 12 2 US-08-928-958-16
106 5 2.0 12 2 US-09-372-429-16
107 5 2.0 12 2 US-08-433-908B-21
108 5 2.0 12 4 US-08-410-614-21
109 5 2.0 12 5 PCI-US95-05471-50
110 5 2.0 13 1 US-08-230-047-19
111 5 2.0 13 2 US-08-572-951-19
112 5 2.0 13 2 US-08-572-951-20
113 5 2.0 13 2 US-08-572-951-40
114 5 2.0 14 1 US-07-745-382-5
115 5 2.0 14 1 US-07-921-848-5
116 5 2.0 14 1 US-07-841-997A-8
117 5 2.0 14 1 US-08-165-301A-5
118 5 2.0 14 1 US-07-958-903A-37
119 5 2.0 14 1 US-07-789-912-6
120 5 2.0 14 1 US-08-321-668-38
121 5 2.0 14 1 US-08-462-078-37
122 5 2.0 14 1 US-08-837-941-38
123 5 2.0 14 1 US-08-823-245-37
124 5 2.0 14 1 US-08-293-362-8
125 5 2.0 14 4 US-08-810-436-5
126 5 2.0 14 4 US-08-602-999A-402
127 5 2.0 14 4 US-09-066-046-37
128 5 2.0 14 4 US-07-963-329A-37
129 5 2.0 14 4 US-09-013-598-8
130 5 2.0 14 4 US-09-500-124-402
131 5 2.0 14 5 PCI-US92-09443A-37
132 5 2.0 14 5 PCI-US94-14179-5
133 5 2.0 14 6 5314872-4
134 5 2.0 14 6 5464943-29
135 5 2.0 15 1 US-08-210-880B-7

```

ALIGNMENTS

```

RESULT 1
US-07-676-647-5
: Sequence 5, Application US/0767664?
: Patent No. 542617
: GENERAL INFORMATION:
: APPLICANT: Davis, Samuel
: APPLICANT: Squinto, Stephen P.
: APPLICANT: Furth, Mark E.
: TITLE OF INVENTION: The Ciliary Neurotrophic Factor Receptor
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennic & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07676,647
: FILING DATE: 19910328
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Mistrock, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 6526-048
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212 790-9090

```

```

: TELEFAX: 212 8698864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 62 amino acids
: TYPE: AMINO ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-07-676-647-5

: Query Match 2.8% Score 7: DB 1: Length 62:
: Best Local Similarity 100.0%; Pred. No. 9.4;
: Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AELSDSG 220
Db 51 AELSDSG 57
|||||

RESULT 2
US-08-449-329-5
: Sequence 5, Application US/08449329
: Patent No. 5648334
: GENERAL INFORMATION:
: APPLICANT: Davis, Samuel
: APPLICANT: Squinto, Stephen P.
: APPLICANT: Furth, Mark E.
: APPLICANT: Yancopoulos, George D.
: TITLE OF INVENTION: The Ciliary Neurotrophic Factor Receptor
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennic & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08449,329
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/700,677
: FILING DATE: 15-MAY-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Mistrock, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 6526-065
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212 790-9090
: TELEFAX: 212 8698864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 62 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
US-08-449-329-5

: Query Match 2.8% Score 7: DB 1: Length 62:
: Best Local Similarity 100.0%; Pred. No. 9.4;
: Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AELSDSG 220
|||||

```

Db 51 AELSDSG 57

RESULT 3

US-08-445-073-5
 : Sequence 5, Application US/08445073
 : Patent No. 5849697
 : GENERAL INFORMATION:
 : APPLICANT: Davis, Samuel
 : APPLICANT: Squinto, Stephen P.
 : APPLICANT: Furth, Mark E.
 : APPLICANT: Yancopoulos, George D.
 : TITLE OF INVENTION: The Ciliary Neurotrophic Factor Receptor
 : NUMBER OF SEQUENCES: 15
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/445,073
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/676,647
 : FILING DATE: 28-MAR-1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Misrock, S. Leslie
 : REGISTRATION NUMBER: 18,872
 : REFERENCE/DOCKET NUMBER: 6526-048
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 212 790-9090
 : TELEFAX: 212 8698864/9741
 : TELEX: 66141 PENNTE
 : INFORMATION FOR SEQ ID NO: 5:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 62 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: linear
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : US-08-445-073-5

Query Match 2.8% Score 7; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 AELSDSG 220

Db 51 AELSDSG 57

RESULT 4

US-08-585-258-5
 : Sequence 5, Application US/08585258
 : Patent No. 5892003
 : GENERAL INFORMATION:
 : APPLICANT: Davis, Samuel
 : APPLICANT: Squinto, Stephen P.
 : APPLICANT: Furth, Mark E.
 : APPLICANT: Yancopoulos, George D.
 : TITLE OF INVENTION: The Ciliary Neurotrophic Factor Receptor
 : NUMBER OF SEQUENCES: 15
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 : STREET: 1155 Avenue of the Americas

CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/585,258
 FILING DATE: 11-JAN-1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/901,904
 FILING DATE:
 APPLICATION NUMBER: US/07/700,677
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 6526-065
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 8698864/9741
 TELEX: 66141 PENNTE
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 62 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-585-258-5

Query Match 2.8% Score 7; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 214 AELSDSG 220

Db 51 AELSDSG 57

RESULT 5

US-09-211-590-5
 : Sequence 5, Application US/09211590
 : Patent No. 6316206
 : GENERAL INFORMATION:
 : APPLICANT: Davis, Samuel
 : APPLICANT: Squinto, Stephen P.
 : APPLICANT: Furth, Mark E.
 : APPLICANT: Yancopoulos, George D.
 : TITLE OF INVENTION: The Ciliary Neurotrophic Factor Receptor
 : NUMBER OF SEQUENCES: 15
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/211,590
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/676,647

1 FILING DATE: 28-MAR-1991
 2 ATTORNEY/AGENT INFORMATION:
 3 NAME: MISTOCK, S. Leslie
 4 REGISTRATION NUMBER: 18,872
 5 REFERENCE/DOCKET NUMBER: 6526-065
 6 TELEPHONE: 212 790-9090
 7 TELEFAX: 212 8698864/9741
 8 TELEX: 66141 PENNIE
 9 INFORMATION FOR SEQ ID NO: 5:
 10 SEQUENCE CHARACTERISTICS:
 11 LENGTH: 62 amino acids
 12 TYPE: amino acid
 13 STRANDEDNESS: single
 14 TOPOLOGY: linear
 15 MOLECULE TYPE: peptide
 16 US-09-211-590-5

Query Match 2.8%; Score 7; DB 4; Length 62;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AELSDSG 220
 DB 51 AELSDSG 57

1 RESULT 5
 2 PCT-US91-03896-5
 3 Sequence 5, Application PCT/US91/03896
 4 GENERAL INFORMATION:
 5 APPLICANT: Davis, Samuel
 6 APPLICANT: Squinto, Stephen P.
 7 APPLICANT: Furth, Mark E.
 8 APPLICANT: Vancopoulos, George D.
 9 TITLE OF INVENTION: The Ciliary Neurotrophic Factor Receptor
 10 NUMBER OF SEQUENCES: 15
 11 CORRESPONDENCE ADDRESS:
 12 ADDRESSEE: Pennie & Edmonds
 13 STREET: 1155 Avenue of the Americas
 14 CITY: New York
 15 STATE: New York
 16 COUNTRY: U.S.A.
 17 ZIP: 10036
 18 COMPUTER READABLE FORM:
 19 MEDIUM TYPE: Floppy disk
 20 COMPUTER: IBM PC compatible
 21 OPERATING SYSTEM: PC-DOS/MS-DOS
 22 SOFTWARE: Patent In Release #1.0, Version #1.25
 23 CURRENT APPLICATION DATA:
 24 APPLICATION NUMBER: PCT/US91/03896
 25 FILING DATE: 19910603
 26 CLASSIFICATION: 435
 27 ATTORNEY/AGENT INFORMATION:
 28 NAME: MISTOCK, S. Leslie
 29 REGISTRATION NUMBER: 18,872
 30 REFERENCE/DOCKET NUMBER: 6526-065-228
 31 TELECOMMUNICATION INFORMATION:
 32 TELEPHONE: 212 790-9090
 33 TELEFAX: 212 8698864/9741
 34 TELEX: 66141 PENNIE
 35 INFORMATION FOR SEQ ID NO: 5:
 36 SEQUENCE CHARACTERISTICS:
 37 LENGTH: 62 amino acids
 38 TYPE: AMINO ACID
 39 STRANDEDNESS: single
 40 TOPOLOGY: unknown
 41 MOLECULE TYPE: peptide
 42 PCT-US91-03896-5

Query Match 2.8%; Score 7; DB 5; Length 62;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QY 214 AELSDSG 220
 2 DB 51 AELSDSG 57
 3
 4 RESULT 7
 5 US-08-672-610A-39
 6 Sequence 39, Application US/08672610A
 7 Patent No. 5846933
 8 GENERAL INFORMATION:
 9 APPLICANT: Ziwel, Hwang
 10 APPLICANT: Korgold, Robert
 11 TITLE OF INVENTION: CD4-DERIVED PEPTIDES THAT INHIBIT IMMUNE
 12 RESPONSES
 13 NUMBER OF SEQUENCES: 53
 14 CORRESPONDENCE ADDRESS:
 15 ADDRESSEE: PENNIE & EDMONDS
 16 STREET: 1155 Avenue of the Americas
 17 CITY: New York
 18 STATE: New York
 19 COUNTRY: United States
 20 ZIP: 10036-2711
 21 COMPUTER READABLE FORM:
 22 MEDIUM TYPE: Floppy disk
 23 COMPUTER: IBM PC compatible
 24 OPERATING SYSTEM: PC-DOS/MS-DOS
 25 SOFTWARE: Patent In Release #1.0, Version #1.30
 26 CURRENT APPLICATION DATA:
 27 APPLICATION NUMBER: US/08/672,610A
 28 FILING DATE:
 29 CLASSIFICATION: 435
 30 ATTORNEY/AGENT INFORMATION:
 31 NAME: FriebeL, Thomas E.
 32 REGISTRATION NUMBER: 29,258
 33 REFERENCE/DOCKET NUMBER: 8666-003
 34 TELECOMMUNICATION INFORMATION:
 35 TELEPHONE: (212) 790-9090
 36 TELEFAX: (212) 869-8864/9741
 37 TELEX: 66141 PENNIE
 38 INFORMATION FOR SEQ ID NO: 39:
 39 SEQUENCE CHARACTERISTICS:
 40 LENGTH: 6 amino acids
 41 TYPE: amino acid
 42 STRANDEDNESS:
 43 TOPOLOGY: unknown
 44 MOLECULE TYPE: peptide
 45 US-08-672-610A-39

Query Match 2.4%; Score 6; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSGQ 221
 DB 1 LSDSGQ 6

1 RESULT 8
 2 US-08-672-610A-40
 3 Sequence 40, Application US/08672610A
 4 Patent No. 5846933
 5 GENERAL INFORMATION:
 6 APPLICANT: Ziwel, Hwang
 7 APPLICANT: Korgold, Robert
 8 TITLE OF INVENTION: CD4-DERIVED PEPTIDES THAT INHIBIT IMMUNE
 9 RESPONSES
 10 NUMBER OF SEQUENCES: 53
 11 CORRESPONDENCE ADDRESS:
 12 ADDRESSEE: PENNIE & EDMONDS
 13 STREET: 1155 Avenue of the Americas
 14 CITY: New York
 15 STATE: New York

COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: E-copy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/672,610A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Friebe, Thomas E.
REGISTRATION NUMBER: 29,258
REFERENCE/DOCKET NUMBER: 8666-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
FLEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-672-610A-40

Query Match 2.4%; Score 6; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSGQ 221
|||||
DB 1 LSDSGQ 6

RESULT 9
US-08-672-610A-6
Sequence 51, Application US/08672610A
Patent No. 5846933
GENERAL INFORMATION:
APPLICANT: Ziwel, Hwang
TITLE OF INVENTION: CD4-DERIVED PEPTIDES THAT INHIBIT IMMUNE
RESPONSES
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/672,610A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Friebe, Thomas E.
REGISTRATION NUMBER: 29,258
REFERENCE/DOCKET NUMBER: 8666-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
FLEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-672-610A-6

Query Match 2.4%; Score 6; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSGQ 221
|||||
DB 1 LSDSGQ 6

RESULT 10
US-08-672-610A-51
Sequence 51, Application US/08672610A
Patent No. 5846933
GENERAL INFORMATION:
APPLICANT: Ziwel, Hwang
TITLE OF INVENTION: CD4-DERIVED PEPTIDES THAT INHIBIT IMMUNE
RESPONSES
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/672,610A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Friebe, Thomas E.
REGISTRATION NUMBER: 29,258
REFERENCE/DOCKET NUMBER: 8666-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
FLEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-672-610A-51

Query Match 2.4%; Score 6; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LSDSGQ 221
|||||
DB 1 LSDSGQ 6

RESULT 11
US-08-257-783-9
Sequence 9, Application US/08257783
Patent No. 6100377
GENERAL INFORMATION:

```

1  APPLICANT: Mark I. Greene
2  TITLE OF INVENTION: CONSTRAINED PEPTIDES
3  NUMBER OF SEQUENCES: 11
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz
6  STREET: One Liberty Place - 46th Floor
7  CITY: Philadelphia
8  STATE: PA
9  COUNTRY: U.S.A.
10 ZIP: 19103
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: 3.5 inch disk, 720 KB
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC DOS/MS-DOS
16 SOFTWARE: WordPerfect 5.1
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/08/257,783
19 FILING DATE: Renewith
20 CLASSIFICATION:
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER:
23 FILING DATE:
24
25 ATTORNEY/AGENT INFORMATION:
26 NAME: John W. Caldwell
27 REGISTRATION NUMBER: 28,937
28 REFERENCE/DOCKET NUMBER: UPN-1657
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 215-568-3100
31 TELEFAX: 215-568-3439
32
33 INFORMATION FOR SEQ ID NO: 9:
34 LENGTH: 12 amino acid residues
35 TYPE: amine acid
36 TOPOLOGY: unknown
37
38 US-08-257-783-9
39
40 Query Match 2.4%, Score 6; DB 3; Length 12;
41 Best Local Similarity 100.0%; Pred. No. 20;
42 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
43
44 QY 216 LSDSGQ 221
45 Db 3 LSDSGQ 8
46
47 RESULT 12
48 PCT-US95-07157-9
49 Sequence 9, Application PC/TUS9507157
50 GENERAL INFORMATION:
51 APPLICANT: Mark I. Greene
52 TITLE OF INVENTION: CONSTRAINED PEPTIDES
53 NUMBER OF SEQUENCES: 11
54 CORRESPONDENCE ADDRESS:
55 ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz
56 STREET: One Liberty Place - 46th Floor
57 CITY: Philadelphia
58 STATE: PA
59 COUNTRY: U.S.A.
60 ZIP: 19103
61
62 COMPUTER READABLE FORM:
63 MEDIUM TYPE: 3.5 inch disk, 720 KB
64 COMPUTER: IBM PC compatible
65 OPERATING SYSTEM: PC-DOS/MS-DOS
66 SOFTWARE: WordPerfect 5.1
67 CURRENT APPLICATION DATA:
68 APPLICATION NUMBER: PCT/US95/07157
69 FILING DATE:
70 CLASSIFICATION:
71 PRIOR APPLICATION DATA:
72 APPLICATION NUMBER: US 08/257,783
73 FILING DATE: 10-30N-1994

```

```

1  ATTORNEY/AGENT INFORMATION:
2  NAME: Mark DeLuca
3  REGISTRATION NUMBER: 33,229
4  REFERENCE/DOCKET NUMBER: UPN-2355
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE: 215-568-3100
7  TELEFAX: 215-568-3439
8  INFORMATION FOR SEQ ID NO: 9:
9  SEQUENCE CHARACTERISTICS:
10 LENGTH: 12 amino acid residues
11 TYPE: amino acid
12 TOPOLOGY: unknown
13
14 PCT-US95-07157-9
15
16 Query Match 2.4%, Score 6; DB 5; Length 12;
17 Best Local Similarity 100.0%; Pred. No. 20;
18 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
19
20 QY 216 LSDSGQ 221
21 Db 3 LSDSGQ 8
22
23 RESULT 13
24 US-08-893-526A-31
25 Sequence 31, Application US/08893526A
26 Patent No. 6096707
27 GENERAL INFORMATION:
28 APPLICANT: Heino, Jyrki
29 APPLICANT: Ivaska, Johanna
30 APPLICANT: K Pyl, Jarmo
31 TITLE OF INVENTION: Integrin Binding Peptide and Use Thereof
32 NUMBER OF SEQUENCES: 38
33 CORRESPONDENCE ADDRESS:
34 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
35 STREET: 1100 New York Ave., N.W.
36 CITY: Washington
37 STATE: D.C.
38 COUNTRY: USA
39 ZIP: 20005-3934
40
41 COMPUTER READABLE FORM:
42 MEDIUM TYPE: Floppy disk
43 COMPUTER: IBM PC compatible
44 OPERATING SYSTEM: PC-DOS/MS-DOS
45 SOFTWARE: PatentIn Release #1.0, Version #1.30
46 CURRENT APPLICATION DATA:
47 APPLICATION NUMBER: US/08/893,526A
48 FILING DATE: 11-JUL-1997
49 CLASSIFICATION: 435
50 ATTORNEY/AGENT INFORMATION:
51 NAME: KIM, JUDITH U.
52 REGISTRATION NUMBER: 40,679
53 REFERENCE/DOCKET NUMBER: 1708.0240000
54 TELECOMMUNICATION INFORMATION:
55 TELEPHONE: 202/371-2600
56 TELEFAX: 202/371-2540
57 TELEX:
58 INFORMATION FOR SEQ ID NO: 31:
59 SEQUENCE CHARACTERISTICS:
60 LENGTH: 16 amino acids
61 TYPE: amino acid
62 TOPOLOGY: linear
63 MOLECULE TYPE: peptide
64
65 US-08-893-526A-31
66
67 Query Match 2.4%, Score 6; DB 3; Length 16;
68 Best Local Similarity 100.0%; Pred. No. 26;
69 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
70
71 QY 189 ALLEKA 194
72 Db 5 ALLEKA 10

```

RESULT 14
US-08-047-033-11
Sequence 11, Application US/08/047033
Patent No. 5444157
GENERAL INFORMATION:
APPLICANT: Fujio SUZUKI et al.
TITLE OF INVENTION: CHONDROMODULIN I PROTEIN
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/047,033
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/821,859
FILING DATE:
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8855
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 7
IDENTIFICATION METHOD:
OTHER INFORMATION: /note- "Xaa is Val or Met"
FEATURE:
NAME/KEY: modified site
LOCATION: 9

IDENTIFICATION METHOD:
OTHER INFORMATION: /note- "Xaa is unknown"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-047-033-11
Query Match 2.4%, Score 6; DB 1; Length 28;
Best local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;
QY 227 GGPARG 232
Db 21 GGPARG 26
RESULT 15
US-08-851-843A-178
Sequence 178, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809c1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00293005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

APPLICATION NUMBER:

APPLICATION NUMBER: US 08/915,503

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,459
FILING DATE: 18-APR-1997
CLASSIFICATION: 536

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 178:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: linear
MOLECULE TYPE: peptide
US-08-854-050-178

Query Match 2.4% Score 6: DB 4: Length 35;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 243 GAUGLA 248
DB 10 GAUGLA 15
|||||

RESULT 18
US-09-430-323-178
Sequence 178, Application US/09430323
Patent No. 6309867

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
Lingner, Joachim
Nakamura, Toru
Chapman, Karen A.
Morin, Greg B.
Harley, Calvin
Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: 170 Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 05-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 178:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 178:
US-09-430-323-178

Query Match 2.4% Score 6: DB 4: Length 35;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 243 GAUGLA 248
DB 10 GAUGLA 15
|||||

RESULT 19
US-08-472-244-1
Sequence 1, Application US/08472244
Patent No. 5821088

GENERAL INFORMATION:
APPLICANT: Darzins, Aldis
APPLICANT: Whitehead, Stephen
APPLICANT: Hruby, Dennis F.
TITLE OF INVENTION: Use of Gram-Positive Bacteria to Express
TITLE OF INVENTION: Recombinant Proteins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWICKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,244
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 016921-065
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FRAGMENT TYPE: C-terminal
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: Protein A
US-08-472-244-1

Query Match 2.4% Score 6: DB 2: Length 43;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

QY 193 KAQALP 198
 Db 5 KAQALP 10

RESULT 20
 US-08-340-428B-36
 : Sequence 36, Application US/08340428B
 : Patent No. 5648465
 : GENERAL INFORMATION:
 : APPLICANT: MARGOLIS, Richard U.
 : APPLICANT: RAUCH, Uwe
 : APPLICANT: MARGOLIS, Renee K.
 : TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
 : NUMBER OF SEQUENCES: 49
 : TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
 : NUMBER OF SEQUENCES: 49
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Browdy and Neimark
 : STREET: 419 Seventh Street, N.W.
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: U.S.A.
 : ZIP: 20004
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/340-428B
 : FILING DATE: 14 No. 5648465ember 1994
 : CLASSIFICATION: 514
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/522-911
 : FILING DATE: 03 August 1992
 : CLASSIFICATION: 514
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Browdy, Roger L.
 : REGISTRATION NUMBER: 25-618
 : REFERENCE/DOCKET NUMBER: MARGOLIS-1A
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 202-628-5197
 : TELEFAX: 202-737-3528
 : INFORMATION FOR SEQ ID NO: 36:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 46 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : US-08-340-428B-36

Query Match 2.48; Score 6; DB 1; Length 46;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 APGEAS 235
 Db 19 APGEAS 24

RESULT 21
 PCT-US93-07306-36
 : Sequence 36, Application PC/TUS9307306
 : GENERAL INFORMATION:
 : APPLICANT: MARGOLIS, Richard U.
 : APPLICANT: RAUCH, Uwe
 : APPLICANT: MARGOLIS, Renee K.
 : TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
 : NUMBER OF SEQUENCES: 48
 : CORRESPONDENCE ADDRESS:

: ADDRESSEE: Browdy and Neimark
 : STREET: 419 Seventh Street, N.W.
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: U.S.A.
 : ZIP: 20004
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PC/TUS93/07306
 : FILING DATE: 03-AUG-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/522,911
 : FILING DATE: 03-AUG-1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Townsend, Guy K.
 : REGISTRATION NUMBER: 34,033
 : REFERENCE/DOCKET NUMBER: Margolis-1A PCT
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 202-628-5197
 : TELEFAX: 202-737-3528
 : INFORMATION FOR SEQ ID NO: 36:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 46 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : PCT-US93-07306-36

Query Match 2.48; Score 6; DB 5; Length 46;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 APGEAS 235
 Db 19 APGEAS 24

RESULT 22
 US-08-294-189-16
 : Sequence 16, Application US/08294189
 : Patent No. 5707822
 : GENERAL INFORMATION:
 : APPLICANT: Fischetti, Vincent A.
 : APPLICANT: Rakonjac, Jasna
 : APPLICANT: Robbins, John
 : TITLE OF INVENTION: GENE SERUM OPACITY FACTOR
 : NUMBER OF SEQUENCES: 23
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Klauber & Jackson
 : STREET: 411 Hackensack Avenue
 : CITY: Hackensack
 : STATE: New Jersey
 : COUNTRY: USA
 : ZIP: 07601
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/294,189
 : FILING DATE: 22-AUG-1994
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Jackson Esq., David A.
 : REGISTRATION NUMBER: 26,742
 : REFERENCE/DOCKET NUMBER: 600-1-118
 : TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800
TELEFAX: 201-343-1584
TELEX: 133521
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYDROTHERMAL: NO
FRAGMENT TYPE: C-terminal
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
US-08-294-189-16

Query Match 2.4%; Score 6; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 79; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 193 ALLEKA 198
|||||
DB 12 ALLEKA 17

RESULT 23
US-09-120-365-94
Sequence 94, Application US/09120365
Patent No. 6103514
GENERAL INFORMATION:
APPLICANT: Natori, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/120.365
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 94
LENGTH: 67
TYPE: PRT
ORGANISM: Homo sapien
US-09-120-365-94

Query Match 2.4%; Score 6; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 11e+02; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 189 ALLEKA 194
|||||
DB 7 ALLEKA 12

RESULT 24
US-09-120-365-95
Sequence 95, Application US/09120365
Patent No. 6103514
GENERAL INFORMATION:
APPLICANT: Natori, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/120.365
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 95
LENGTH: 67
TYPE: PRT
ORGANISM: rattus
US-09-120-365-95

Query Match 2.4%; Score 6; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 189 ALLEKA 194
|||||
DB 7 ALLEKA 12

RESULT 25
US-09-120-365-96
Sequence 96, Application US/09120365
Patent No. 6103514
GENERAL INFORMATION:
APPLICANT: Natori, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/120.365
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 96
LENGTH: 67
TYPE: PRT
ORGANISM: chicken
US-09-120-365-96

Query Match 2.4%; Score 6; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 189 ALLEKA 194
|||||
DB 7 ALLEKA 12

RESULT 26
US-09-120-365-97
Sequence 97, Application US/09120365
Patent No. 6103514
GENERAL INFORMATION:
APPLICANT: Natori, Shunji
TITLE OF INVENTION: NEW PROTEASE
FILE REFERENCE: 32290-144749
CURRENT APPLICATION NUMBER: US/09/120.365
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: JP 9-333 474
EARLIER FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 97
LENGTH: 67
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-120-365-97

Query Match 2.4%; Score 6; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0;
Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 189 ALLEKA 194
|||||
DB 7 ALLEKA 12

RESULT 27
US-09-120-365-98
Sequence 98, Application US/09120365
Patent No. 6103514
GENERAL INFORMATION:

```

; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120-365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 67
; TYPE: PRT
; ORGANISM: rematode
US-09-120-365-98

```

```

Query Match      2.4%; Score 6; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

```

```

QY 189 ALLEKA 194
Db 7 ALLEKA 12

```

```

RESULT 28
US-09-120-365-99
; Sequence 99, Application US/99120365
; Patent No. 6103514
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/120-365
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
; LENGTH: 67
; TYPE: PRT
; ORGANISM: fuke
US-09-120-365-99

```

```

Query Match      2.4%; Score 6; DB 3; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

```

```

QY 189 ALLEKA 194
Db 7 ALLEKA 12

```

```

RESULT 29
US-09-515-039-94
; Sequence 94, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/515,039
; CURRENT FILING DATE: 2000-03-06
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 94
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-515-039-94

```

```

Query Match      2.4%; Score 6; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

```

```

QY 189 ALLEKA 194
Db 7 ALLEKA 12

```

```

RESULT 30
US-09-515-039-95
; Sequence 95, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/515,039
; CURRENT FILING DATE: 2000-03-06
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 67
; TYPE: PRT
; ORGANISM: rattus
US-09-515-039-95

```

```

Query Match      2.4%; Score 6; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

```

```

QY 189 ALLEKA 194
Db 7 ALLEKA 12

```

```

RESULT 31
US-09-515-039-96
; Sequence 96, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NEW PROTEASE
; FILE REFERENCE: 32290-144749
; CURRENT APPLICATION NUMBER: US/09/515,039
; CURRENT FILING DATE: 2000-03-06
; EARLIER APPLICATION NUMBER: JP 9-333 474
; EARLIER FILING DATE: 1997-11-18
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 96
; LENGTH: 67
; TYPE: PRT
; ORGANISM: chicken
US-09-515-039-96

```

```

Query Match      2.4%; Score 6; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

```

```

QY 189 ALLEKA 194
Db 7 ALLEKA 12

```

```

RESULT 32
US-09-515-039-97
; Sequence 97, Application US/09515039
; Patent No. 6214599
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji

```

; TITLE OF INVENTION: NEW PROTEASE
 ; FILE REFERENCE: 32290-144749
 ; CURRENT APPLICATION NUMBER: US/09/515,039
 ; CURRENT FILING DATE: 2000-03-06
 ; EARLIER APPLICATION NUMBER: JP 9-333 474
 ; EARLIER FILING DATE: 1997-11-18
 ; NUMBER OF SEQ ID NOS: 101
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 97
 ; LENGTH: 67
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US 09-515-039 97

Query Match 2.4% Score 6; DB 4; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1.1e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 ALLEKA 194
 Db 7 ALLEKA 12

RESULT 33
 US-09-515-039-99
 ; Sequence 98, Application US/09515039
 ; Patent No. 6214599
 ; GENERAL INFORMATION:
 ; APPLICANT: NALOP, Shoji
 ; TITLE OF INVENTION: NEW PROTEASE
 ; FILE REFERENCE: 32290-144749
 ; CURRENT APPLICATION NUMBER: US/09/515,039
 ; CURRENT FILING DATE: 2000-03-06
 ; EARLIER APPLICATION NUMBER: JP 9-333 474
 ; EARLIER FILING DATE: 1997-11-18
 ; NUMBER OF SEQ ID NOS: 101
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 98
 ; LENGTH: 67
 ; TYPE: PRT
 ; ORGANISM: nematode
 US-09-515-039-99

Query Match 2.4% Score 6; DB 4; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1.1e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 ALLEKA 194
 Db 7 ALLEKA 12

RESULT 34
 US-09-515-039-99
 ; Sequence 99, Application US/09515039
 ; Patent No. 6214599
 ; GENERAL INFORMATION:
 ; APPLICANT: NALOP, Shoji
 ; TITLE OF INVENTION: NEW PROTEASE
 ; FILE REFERENCE: 32290-144749
 ; CURRENT APPLICATION NUMBER: US/09/515,039
 ; CURRENT FILING DATE: 2000-03-06
 ; EARLIER APPLICATION NUMBER: JP 9-333 474
 ; EARLIER FILING DATE: 1997-11-18
 ; NUMBER OF SEQ ID NOS: 101
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 99
 ; LENGTH: 67
 ; TYPE: PRT
 ; ORGANISM: Iluke
 US-09-515-039-99

Query Match 2.4% Score 6; DB 4; Length 67;

Best Local Similarity 100.0%; Pred. No. 1.1e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 189 ALLEKA 194
 Db 7 ALLEKA 12

RESULT 35
 US-09-134-001C-4974
 ; Sequence 4974, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; FILE REFERENCE: GIC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 4974
 ; LENGTH: 70
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-4974

Query Match 2.4% Score 6; DB 4; Length 70;
 Best Local Similarity 100.0%; Pred. No. 1.1e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 SKSILL 101
 Db 60 SKSILL 65

RESULT 36
 US-08-242-097-8
 ; Sequence 8, Application US/08242097
 ; Patent No. 5846763
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, Tae Ho
 ; APPLICANT: Wisniewski, Hans Georg
 ; APPLICANT: Vilcek, Jan
 ; TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Browdy and Neimark
 ; STREET: 419 Seventh Street, N.W., Suite 300
 ; CITY: Washington
 ; STATE: D.C.
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/242,097
 ; FILING DATE: 13-MAY-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BROWDY, Roger L.
 ; REGISTRATION NUMBER: 25,618
 ; REFERENCE/DOCKET NUMBER: LEE26/VILCEK-1B
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-628-5197
 ; TELEFAX: 212-737-3528
 ; INFORMATION FOR SEQ ID NO: 8;

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 5 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-242-097-8

Query Match      2.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SLEPG 11
DB 1 SLEPG 5

RESULT 37
US-08-672-610A-36
: Sequence 36, Application US/08672610A
: Patent No. 5845933
: GENERAL INFORMATION:
: APPLICANT: Ziwel, Hwang
: TITLE OF INVENTION: C24-DERIVED PEPTIDES THAT INHIBIT IMMUNE
: RESPONSES
: NUMBER OF SEQUENCES: 53
: CORRESPONDENCE ADDRESS:
: ADDRESSER: PENNIE & EDMONDS
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/672,610A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Priebe, Thomas E.
: REGISTRATION NUMBER: 29,258
: REFERENCE/DOCKET NUMBER: 8666-003
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 36:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
US-08-672-610A-36

Query Match      2.0%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 SSSGQ 221
DB 1 SSSGQ 5

RESULT 38
US-08-582-740-41
: Sequence 41, Application US/08582740
: Patent No. 6037324
: GENERAL INFORMATION:
: APPLICANT: Schwender, Charles F.
: APPLICANT: Shroff, Hitesh N.
: TITLE OF INVENTION: Inhibitors of MAGCAM-1-Mediated
: TITLE OF INVENTION: Interactions and Methods of Use Therefor
: NUMBER OF SEQUENCES: 70
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Militia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/582,740
: FILING DATE: 04-JAN-1996
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Brook, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: LKS95-12
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 861-6240
: TELEFAX: (617) 861-9540
: INFORMATION FOR SEQ ID NO: 41:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 1
: OTHER INFORMATION: /label= modified aa
: OTHER INFORMATION: /note= "Pda - Leucine"
: FEATURE:
: NAME/KEY: Modified-site
: LOCATION: 5
: OTHER INFORMATION: /label= modified aa
: OTHER INFORMATION: /note= "Leucine - NH2"
US-08-582-740-41

Query Match      2.0%; Score 5; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LATSL 8
DB 1 LATSL 5

RESULT 39
US-09-206-695-8
: Sequence 8, Application US/09206695
: Patent No. 6210905
: GENERAL INFORMATION:
: APPLICANT: Lee, Iac Ho
: APPLICANT: Wisniewski, Hans Georg
: APPLICANT: Vilcek, Jan
: TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
: TITLE OF INVENTION: Therefor and Uses Thereof
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Browdy and Neimark
: STREET: 419 Seventh Street, N.W., Suite 300
: CITY: Washington
: STATE: D.C.

```

```

1  ZIP: 20004
2  COMPUTER READABLE FORM:
3  MEDIUM TYPE: Floppy disk
4  COMPUTER: IBM PC compatible
5  OPERATING SYSTEM: PC-DOS/MS-DOS
6  SOFTWARE: Patent In Release #1.0, Version #1.25
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/09/206,695
9  FILING DATE:
10 CLASSIFICATION:
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US/08/242,097
13 FILING DATE: 13-MAY-1994
14 ATTORNEY/AGENT INFORMATION:
15 NAME: BROWDY, Roger L.
16 REGISTRATION NUMBER: 25,618
17 REFERENCE/DOCKET NUMBER: LEE26/WILCEK-1B
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: 212-628-5197
20 TELEFAX: 212-737-9528
21 INFORMATION FOR SEQ ID NO: 8:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 5 amino acids
24 TYPE: amino acid
25 STRANDEDNESS: single
26 TOPOLOGY: linear
27 MOLECULE TYPE: peptide
28 US-09 206-695-8
29
30 Query Match 2.0%; Score 5; DB 4; Length 5;
31 Best Local Similarity 100.0%; Pred. No. 2e-05;
32 Matches 5; Conservative C; Mismatches 0; Indels 0; Gaps 0;
33
34 QY 7 SLIPS 11
35 1111
36 1 SLIPS 5
37
38 RESULT: 40
39 US-09-109-879-41
40 Sequence 41, Application US/09/109879
41 Patent No. 6274556
42 GENERAL INFORMATION:
43 APPLICANT: Schwender, Charles F.
44 APPLICANT: Shroff, Hitesh N.
45 TITLE OF INVENTION: INHIBITORS OF MACAM-1-MEDIATED
46 TITLE OF INVENTION: INTERACTIONS AND METHODS OF USE THEREFOR
47 NUMBER OF SEQUENCES: 89
48 CORRESPONDENCE ADDRESS:
49 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
50 STREET: Two Militau Drive
51 CITY: Lexington
52 STATE: Massachusetts
53 COUNTRY: USA
54 ZIP: 02421
55 COMPUTER READABLE FORM:
56 MEDIUM TYPE: Floppy disk
57 COMPUTER: IBM PC compatible
58 OPERATING SYSTEM: PC-DOS/MS-DOS
59 SOFTWARE: Patent In Release #1.0, Version #1.30
60 CURRENT APPLICATION DATA:
61 APPLICATION NUMBER: US/09/109,879
62 FILING DATE: 02-JUL-1998
63 CLASSIFICATION: 514
64 PRIOR APPLICATION DATA:
65 APPLICATION NUMBER: PCT/US97/00291
66 FILING DATE: 03-JAN-1997
67 PRIOR APPLICATION DATA:
68 APPLICATION NUMBER: US 08/582,740
69 FILING DATE: 04-JAN-1996
70 ATTORNEY/AGENT INFORMATION:
71 NAME: Brook, David E.
72 REGISTRATION NUMBER: 22,592

```

```

1  REFERENCE/DOCKET NUMBER: LKS95-12A2
2  TELECOMMUNICATION INFORMATION:
3  TELEPHONE: (781) 861-6240
4  TELEFAX: (781) 861-9540
5  INFORMATION FOR SEQ ID NO: 41:
6  SEQUENCE CHARACTERISTICS:
7  LENGTH: 5 amino acids
8  TYPE: amino acid
9  STRANDEDNESS:
10 TOPOLOGY: linear
11 MOLECULE TYPE: peptide
12 FEATURE:
13 NAME/KEY: Modified-site
14 LOCATION: 1
15 OTHER INFORMATION: /label= modified aa
16 OTHER INFORMATION: /note= *pba - Leucine*
17 FEATURE:
18 NAME/KEY: Modified-site
19 LOCATION: 5
20 OTHER INFORMATION: /label= modified aa
21 OTHER INFORMATION: /note= *leucine - NH2*
22 US-09-109-879-41

```

```

1  Query Match 2.0%; Score 5; DB 4; Length 5;
2  Best Local Similarity 100.0%; Pred. No. 2e-05;
3  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
4
5  QY 4 LATS1, 8
6  11111
7  Db 1 LATS1 5

```

```

1  Search completed: April 10, 2003, 10:42:55
2  Job time : 13.2941 secs

```


GenCore version 5.1.4_p5.457e
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:30:25 : Search time 21.4706 Seconds
(without alignments)
2399.179 Million cell updates/sec

Title: US-09-930-026

Perfect score: 250

Sequence: 1 MALLAISLPNIMVKTFEIR.....PGFASGSSRSGCAQCIAS 250

Scoring table: GL30

Gapep 50.0, Gapext 60.0

Searched: 67:580 seqs, 206347115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 60947

Minimum DB seq length: 3

Maximum DB seq length: 70

Post-processing: listing first 135 summaries

Database: SPTMHBL21:

- 1: sp.archaea:
- 2: sp.bacteria:
- 3: sp.fungi:
- 4: sp.human:
- 5: sp.invertebrate:
- 6: sp.mammal:
- 7: sp.mhc:
- 8: sp.organelle:
- 9: sp.phage:
- 10: sp.plant:
- 11: sp.rodent:
- 12: sp.virus:
- 13: sp.vertebrate:
- 14: sp.unclassified:
- 15: sp.virus:
- 16: sp.bacterioph:
- 17: sp.archaeap:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	2.4	36	Q93MN1	Q93mcl pseudomonas
2	6	2.4	40	Q68343	Q68343 hepatitis c
3	6	2.4	42	Q32648	Q32648 pinus thunb
4	6	2.4	46	Q9FSV7	Q9fsvl fagus sylv
5	6	2.4	48	Q9W066	Q9w066 drosophila
6	6	2.4	54	Q54605	Q54605 streptococ
7	6	2.4	60	Q8VAP2	Q8vap2 white spot
8	6	2.4	63	Q9C1H2	Q9c1h2 lactococcus
9	6	2.4	65	Q93K32	Q93k32 streptococ
10	6	2.4	66	Q9KKF8	Q9kkf8 streptomyce
11	6	2.4	66	Q9KKF7	Q9kkf7 streptomyce
12	6	2.4	68	Q9N095	Q9n095 macaca fasc
13	6	2.4	70	Q91US3	Q91us3 ponline her
14	5	2.0	8	Q9QV15	Q9qv15 rattus sp.
15	5	2.0	11	Q94785	Q94785 homo sapien
16	5	2.0	16	Q92444	Q92444 nicotiana t

17	5	2.0	19	15	Q905J6	Q905j6 human immun
18	5	2.0	21	10	Q9AUF0	Q9auf0 brassica na
19	5	2.0	23	2	Q9R5R3	Q9r5r3 photobacter
20	5	2.0	24	2	P82997	P82997 pseudomonas
21	5	2.0	25	10	Q49748	Q49748 arabidopsis
22	5	2.0	25	10	Q9AUE5	Q9aue5 brassica ca
23	5	2.0	27	4	Q9UP08	Q9up08 homo sapien
24	5	2.0	27	8	Q9T8S6	Q9t8s6 liolaemus c
25	5	2.0	27	8	Q9T8S0	Q9t8s0 liolaemus b
26	5	2.0	27	8	Q9T8R8	Q9t8r8 liolaemus c
27	5	2.0	27	11	Q9QV51	Q9qv51 mus sp. leu
28	5	2.0	27	12	Q9J5Z4	Q9j5z4 hepatitis c
29	5	2.0	27	12	Q9J5Z3	Q9j5z3 hepatitis c
30	5	2.0	27	12	Q9J5Z1	Q9j5z1 hepatitis c
31	5	2.0	27	12	Q9J5Z0	Q9j5z0 hepatitis c
32	5	2.0	27	12	Q9J5Y9	Q9j5y9 hepatitis c
33	5	2.0	27	12	Q9J5Y8	Q9j5y8 hepatitis c
34	5	2.0	27	12	Q9J5Y7	Q9j5y7 hepatitis c
35	5	2.0	27	12	Q9IKT6	Q9ikt6 hepatitis c
36	5	2.0	27	12	Q9IKT4	Q9ikt4 hepatitis c
37	5	2.0	27	12	Q9IKT3	Q9ikt3 hepatitis c
38	5	2.0	27	15	Q9IKX2	Q9ikx2 simian retr
39	5	2.0	28	6	Q9N0N9	Q9n0n9 cercopithe
40	5	2.0	29	2	Q93IG2	Q93ig2 thiobacillu
41	5	2.0	29	6	P79112	P79112 bos taurus
42	5	2.0	30	2	P83001	P83001 pseudomonas
43	5	2.0	30	3	Q9URA7	Q9ura7 fusarium ox
44	5	2.0	30	5	Q9BW72	Q9bm72 drosophila
45	5	2.0	30	6	Q9TQ01	Q9tq01 equus cabal
46	5	2.0	30	10	Q41490	Q41490 solanum tub
47	5	2.0	30	15	Q9WJA9	Q9wj49 human immun
48	5	2.0	31	12	O56694	O56694 hepatitis c
49	5	2.0	32	6	Q9GL62	Q9gl62 bos taurus
50	5	2.0	32	10	Q9S9C3	Q9s9c3 arabidopsis
51	5	2.0	33	4	Q9UE40	Q9ue40 homo sapien
52	5	2.0	33	13	Q8U0R8	Q8uur8 brachydanio
53	5	2.0	34	3	Q94179	Q94179 neurospora
54	5	2.0	35	2	Q9R5T9	Q9r5t9 campylobact
55	5	2.0	35	4	Q94783	Q94783 homo sapien
56	5	2.0	35	10	Q9FT84	Q9ft84 arabidopsis
57	5	2.0	35	10	Q9Z0W2	Q9z0w2 arabidopsis
58	5	2.0	35	11	Q9UJ92	Q9uju2 mus musculu
59	5	2.0	36	5	Q9IXF0	Q9ixf0 osteritagia
60	5	2.0	36	10	Q9LD51	Q9ld51 psilotam nu
61	5	2.0	37	2	Q60240	Q60240 pseudomonas
62	5	2.0	37	15	Q85487	Q85487 bovine leuk
63	5	2.0	37	15	Q57050	Q57050 human immun
64	5	2.0	38	2	Q68802	Q68802 pseudomonas
65	5	2.0	38	2	Q9REY8	Q9rey8 carnobacter
66	5	2.0	38	11	Q92500	Q92500 cricetus
67	5	2.0	38	15	Q88391	Q88391 simian t-ly
68	5	2.0	38	15	Q88405	Q88405 simian t-ly
69	5	2.0	38	15	Q88409	Q88409 simian t-ly
70	5	2.0	38	15	Q88410	Q88410 simian t-ly
71	5	2.0	38	15	Q88411	Q88411 simian t-ly
72	5	2.0	38	15	Q88393	Q88393 simian t-ly
73	5	2.0	38	15	Q88395	Q88395 simian t-ly
74	5	2.0	38	15	Q88397	Q88397 simian t-ly
75	5	2.0	38	15	Q80788	Q80788 human t-lym
76	5	2.0	38	15	Q88355	Q88355 simian t-ly
77	5	2.0	38	15	Q88361	Q88361 simian t-ly
78	5	2.0	38	15	Q88367	Q88367 simian t-ly
79	5	2.0	38	15	Q88372	Q88372 simian t-ly
80	5	2.0	38	15	Q88378	Q88378 simian t-ly
81	5	2.0	38	15	Q88382	Q88382 simian t-ly
82	5	2.0	38	15	Q80816	Q80816 human t-lym
83	5	2.0	38	15	Q12811	Q12811 simian t-ly
84	5	2.0	38	15	Q12812	Q12812 simian t-ly
85	5	2.0	38	15	P80649	P80649 human t-cell
86	5	2.0	38	15	P90346	P90346 human t-cell
87	5	2.0	38	15	P96651	P96651 human t-lym
88	5	2.0	38	15	Q12813	Q12813 simian t-ly
89	5	2.0	38	15	Q12814	Q12814 simian t-ly

90 5 2.0 36 15 012815
 91 5 2.0 38 15 012816
 92 5 2.0 38 15 P89674
 93 5 2.0 38 15 P90459
 94 5 2.0 39 4 Q90QV0
 95 5 2.0 39 4 016561
 96 5 2.0 39 12 Q83757
 97 5 2.0 39 15 P89650
 98 5 2.0 39 15 Q85092
 99 5 2.0 39 15 Q9YM63
 100 5 2.0 39 16 Q9PC47
 101 5 2.0 39 16 Q8U5X2
 102 5 2.0 39 16 Q8U5S7
 103 5 2.0 40 4 Q9UE36
 104 5 2.0 40 5 Q9N7S7
 105 5 2.0 40 8 Q9U533
 106 5 2.0 40 10 Q9AUE8
 107 5 2.0 40 10 Q94E10
 108 5 2.0 40 11 Q8K112
 109 5 2.0 40 12 Q91K67
 110 5 2.0 40 12 Q91K54
 111 5 2.0 40 12 Q91K34
 112 5 2.0 40 12 Q91J26
 113 5 2.0 40 15 Q07620
 114 5 2.0 40 16 Q8VJY5
 115 5 2.0 41 5 Q9DKS5
 116 5 2.0 41 9 Q91136
 117 5 2.0 41 10 Q9AUE6
 118 5 2.0 41 16 Q98FE3
 119 5 2.0 42 15 Q82214
 120 5 2.0 42 15 Q82215
 121 5 2.0 42 16 Q9WYX4
 122 5 2.0 43 12 Q9E6H5
 123 5 2.0 45 10 Q41620
 124 5 2.0 45 16 Q8U5F2
 125 5 2.0 46 2 Q87100
 126 5 2.0 46 2 Q82943
 127 5 2.0 46 2 Q9F1F0
 128 5 2.0 46 2 Q9Z1D2
 129 5 2.0 46 9 Q9MWT1
 130 5 2.0 46 15 Q80812
 131 5 2.0 47 2 Q54871
 132 5 2.0 47 2 Q8VRK1
 133 5 2.0 47 5 Q9VE02
 134 5 2.0 47 6 Q46565
 135 5 2.0 47 6 Q9TQX9

ALIGNMENTS

RESULTS 1

Q93MN: PRELIMINARY: PRI: 38 AA.
 ID Q93MN: PRELIMINARY: PRI: 38 AA.
 AC Q93MN: PRELIMINARY: PRI: 38 AA.
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Histidine protein kinase sensor CorS (Fragment).
 OS Pseudomonas syringae (pv. actinidiae).
 GC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 NCBI_TaxID=103796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jung J.S., Han H.S., Koh Y.J.;
 RT "Production of the phytotoxin coronatine in Pseudomonas syringae pv.
 RT actinidiae isolated from Korea";
 RI Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
 DR FMBI; AF450252; AAK94469.1;
 KW Kinase.
 ET NON_TER 38 38
 SQ SEQUENCE 36 AA: 4433 MW: 442680AFC7913 CRC64;

012815 simian t-ly
 012816 simian t-ly
 P89674 simian t-ly
 P90459 simian t-ly
 Q90QV0 homo sapien
 Q16561 homo sapien
 Q83757 norwalk vir
 P89650 human t-cel
 Q85092 primate t-1
 Q9YM63 human immun
 Q9PC47 xylella fas
 Q8U5X2 agrobacteri
 Q8U5S7 agrobacteri
 Q9UE36 homo sapien
 Q9N7S7 leishmania
 Q9U533 petunia sp.
 Q9AUE8 brassica na
 Q94E10 quassipolide
 Q8112 mus musculu
 Q91K67 hepatitis c
 Q91K54 hepatitis c
 Q91K34 hepatitis c
 Q91J26 hepatitis c
 Q07620 friend sple
 Q8VJY5 mycobacteri
 Q9DKS5 caenorhabdi
 Q91136 bacterioph
 Q9AUE6 brassica ol
 Q98FE3 rhizobium l
 Q82214 human t-cel
 Q82215 human t-cel
 Q9WYX4 thermotoga
 Q9E6H5 dengue viru
 Q41620 trifolium s
 Q8U5F2 agrobacteri
 Q87100 bacillus su
 Q82943 chromatium
 Q9F1F0 enterococcu
 Q9Z1D2 mycoplasma
 Q9MWT1 staphylococ
 Q80812 human t-lym
 Q54871 streptococ
 Q8VRK1 burkholderi
 Q9VE02 drosophila
 Q46565 dama dama
 Q9TQX9 bos taurus

Query Match 2.4%; Score 6; DB 2; Length 38;
 Best Local Similarity 100.0%; Pred.No.1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 3 LLATSL 8
 Db 28 LLATSL 33

RESULTS 2

Q68343 PRELIMINARY: PRI: 40 AA.
 ID Q68343: PRELIMINARY: PRI: 40 AA.
 AC Q68343: PRELIMINARY: PRI: 40 AA.
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1) (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCHI_TaxID=111103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95395020; PubMed=7665662;
 RX Odeberg J., Yun Z., Sonnerborg A., Uhlen M., Lundberg J.;
 R "Dynamic analysis of heterogeneous hepatitis C virus populations by
 R direct solid-phase sequencing";
 RJ J. Clin. Microbiol. 33:1870-1874(1995).
 RL EMBL; 024622; AAA89056.1;
 DR InterPro: IPR002531; HCV_NS1.
 DR Pfam: PF01560; HCV_NS1.
 DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1 1
 FT NON_TER 40 40
 SQ SEQUENCE 40 AA: 3967 MW: 83887EE1975B884 CRC64;

Query Match 2.4%; Score 6; DB 12; Length 40;
 Best Local Similarity 100.0%; Pred.No.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 243 GAQGLA 248
 Db 19 GAQGLA 24

RESULTS 3

Q32948 PRELIMINARY: PRI: 42 AA.
 ID Q32948: PRELIMINARY: PRI: 42 AA.
 AC Q32948: PRELIMINARY: PRI: 42 AA.
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ORF42a.
 OS Pinus thunbergii (Green pine) (Japanese black pine).
 OG Chloroplast.
 OC Fukuyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 NCBI_TaxID=3350;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9221283; PubMed=1557027;
 RA Tsudzuki J., Nakashima K., Tsudzuki T., Hiratsuka J., Shibata M.,
 RA Waksugi T., Suglura M.;
 R "Chloroplast DNA of black pine retains a residual inverted repeat
 R lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and
 R trnH and the absence of rps16";
 RI Mol. Gen. Genet. 232:206-214(1992).
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95094312; PubMed=8001170;

RA Tsudzuki J., Ito S., Tsudzuki I., Wakasugi T., Sugiura M.;
 RI "A new gene encoding tRNA pro (CGG) is present in the chloroplast
 R genome of black pine: a compilation of 42 tRNA genes from black pine
 RT chloroplasts.";
 RL Curr. Genet. 26:153-158(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95024047; PubMed-7937893;
 RA Wakasugi I., Tsudzuki J., Ito S., Nakashima K., Tsudzuki I.,
 RA Sugiura M.;
 RI "Loss of all tRNA genes as determined by sequencing the entire
 RT chloroplast genome of the black pine pines thunbergii.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
 RW EMBL: D17513; HAA:4352.1; -;
 KW Chloroplast.
 SQ SEQUENCE 42 AA; 4972 MW; 9B06BF461A3F8565 CRC64;
 Query Match 2.4%; Score 6; DB 8; Length 42;
 Best Local Similarity 100.0%; Pred. No. 2e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 66 GR-NNP 71
 DB 23 GR-NNP 28
 RESULT 4
 ID 09FSV1 PRELIMINARY; PRI; 45 AA.
 AC 09FSV1;
 DT 01-MAR-2001 (TREMBLrel. 15, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Protein kinase (Fragment).
 GN PKF4.
 OS Fagus sylvatica (Hecchnut).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidaceae;
 CC eurosids I; Fagales; Fagaceae; Fagus.
 CX NCBI_TaxID-28930;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX ITSSUP-SRED;
 RA Lorenzo G.;
 RL Thesis (2002), Fisiologia Vegetal, Facultad de Biologia,
 RL Universidad de Salamanca.
 DR EMBL: AJ246942; CAC09570.1; -;
 DR InterPro: IPR000719; Euk.pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR ProDom: PD000020; Euk.pkinase.1.
 DR PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
 KW Kinase.
 FT NON_TER 1
 FT NON_TER 46
 SQ SEQUENCE 46 AA; 5051 MW; C64FEF5748A98922 CRC64;
 Query Match 2.4%; Score 6; DB 10; Length 46;
 Best Local Similarity 100.0%; Pred. No. 2.2e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 96 SKSSL 101
 DB 25 SKSSL 30
 RESULT 5
 ID QW066 PRELIMINARY; PRI; 48 AA.
 AC QW066;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CG13938 protein.

GN CG13938.
 GS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID-7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-BERKELEY;
 RC MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Schlerer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champé M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.K., Miklos G.L.G.,
 RA Allik J.F., Agbayani A., An H.-J., Andrews-Flannock C., Baldwin D.,
 RA Batlle R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman R.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bock J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Golbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ihegwam C.,
 RA Jajall M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostreli A.,
 RA Mounl S.M., Moy M., Murphy B., Murphy T., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacib J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen R.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu B., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RI "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003473; AAF47592.1; -;
 DR FlyBase: FBgn0040692; CG13938.
 SQ SEQUENCE 48 AA; 5005 MW; 9988F285A92F1AEE CRC64;
 Query Match 2.4%; Score 6; DB 5; Length 48;
 Best Local Similarity 100.0%; Pred. No. 2.3e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 226 DGGPAP 231
 DB 42 DGGPAP 47
 RESULT 6
 ID Q54605 PRELIMINARY; PRI; 54 AA.
 AC Q54605;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Emn1 protein (Fragment).
 GN EMBL.
 OS Streptococcus pyogenes.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 CC Streptococcaceae; Streptococcus.

```

OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TAIWAN:
RA LO C.-F., Kou G.-H.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RA Whatore A.M.;
RT "Sequence Analysis of the Emm-Like Gene Family of Streptococcus
RT pyogenes."
RL Thesis (1993), Microbiology, University of Newcastle upon Tyne.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN M TYPE 80;
RX MEDLINE=95138537; PubMed=2621551;
KA Whitmore A.M., Kapur V., Sullivan D.J., Musser J.M., Kehoe M.A.;
RT "Non-concurrent relationships between variation in emm gene sequences
RT and the population genetic structure of group A streptococci."
RL Mol. Microbiol. 14:619-631(1994).
OR EMBL: U12004; AAA94620.1;
RT NON-TER
RT NON-TER
SQ SEQUENCE 54 AA; 58.2 MW; D574F0F09FEEDF8 CRC64;

Query Match 2.4%; Score 6; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.6e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 ISKSSL 100
Db 43 ISKSSL 48

RESULT 7
QSVAP2
ID QSVAP2 PRELIMINARY; PRT: 60 AA.
AC QSVAP2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Wsv359 (WSSV418).
OS White spot syndrome virus (WSSV).
OC Viruses; unclassified viruses.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=25548311; PubMed=11689662;
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT "Complete genome sequence of the shrimp white spot bacilliform
RT virus."
RL J. Virol. 75:11811-11820(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RC Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-TAIWAN:
RX MEDLINE=20517548; PubMed=11062040;
KA Tsai M.F., Yu H.T., Tzeng H.F., Lea J.H., Chou C.M., Huang C.J.,
KA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
RT "Identification and characterization of a shrimp white spot syndrome
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of
RT cellular-type thymidine kinase and thymidylate kinase."
RL Virology 277:110-110(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-TAIWAN:
RX MEDLINE=21844071; PubMed=11853398;
RA Chen L.L., Liu G.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
RA Lo C.F., Kou G.H.;
RT "Identification of a nucleocapsid protein (VP15) gene of shrimp white
RT spot syndrome virus and characterization of the motif important for
RT targeting VP15 to the nuclei of transfected insect cells."
RL Virology 293:44-53(2002).
RN [5]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-TAIWAN:
RA LO C.-F., Kou G.-H.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF332093; AAL33361.1;
DR EMBL: AF440570; AAC89286.1;
SQ SEQUENCE 60 AA; 6377 MW; 9D6FB8FF17FFB61A CRC64;

Query Match 2.4%; Score 6; DB 12; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.9e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 SSPEPP 211
Db 4 SSPEPP 9

RESULT 8
Q9CIH2 PRELIMINARY; PRT: 63 AA.
AC Q9CIH2;
DI 01-JUN-2001 (TrEMBLrel. 17, Created)
DI 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DI 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown protein.
GN YDJB OR L0388.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21235186; PubMed=11337472;
RA Bolotin A., Mincker P., Mauger S., Jaillon O., Maizarme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL: AF006276; AAK04486.1;
KW Complete proteome.
SQ SEQUENCE 63 AA; 7312 MW; FB3D73F6E176SFA7 CRC64;

Query Match 2.4%; Score 6; DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 3e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 167 RQHST 172
Db 10 RQHST 15

RESULT 9
Q93R32 PRELIMINARY; PRT: 65 AA.
AC Q93R32;
DI 01-DEC-2001 (TrEMBLrel. 19, Created)
DI 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative permease of ABC transporter (Fragment).
GN ORF2.
OS Streptococcus criceti.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RA Tamura H., Kato H.;
RT "Identification of an antigen 1/1 homologous pseudogene in
RT Streptococcus cricetus F49."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB067778; BAB62755.1;
FT NON-TER

```

```

SQ SEQUENCE 65 AA: 7565 MW: 42PEAF6F6313BCKF38 CRC64;
Query Match 2.4%; Score 6; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LFSALI 28
DB 38 LFSALI 43

RESULT 10
Q9KKF9
ID Q9KKF8 PRELIMINARY: PRT; 66 AA.
AC Q9KKF8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DI 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
GN TrpB.
OS Streptomyces sp. ASF13.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
CX NCBI_TaxID:63846;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ASF13.
RA Egan S., Weiner P., Kallifidas D., Wellington E.M.H.;
RT "Analysis of streptomycin biosynthetic gene clusters in streptomycetes
RT isolated from soil.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF064198; AAF75105.1;
DR HSSP: P00933; 2TYS
DR InterPro: IPR001926; B6_enzyme_beta.
DR Pfam: PF00291; PALP; 1.
FI NON_TER 1
FI NON_TER 66
SQ SEQUENCE 66 AA: 6908 MW: 97EB6A1C198CF618 CRC64;

Query Match 2.4%; Score 6; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 GLILVN 114
DB 50 GLILVN 55

RESULT 11
Q9KKF7
ID Q9KKF7 PRELIMINARY: PRT; 66 AA.
AC Q9KKF7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DI 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
GN TrpB.
OS Streptomyces sp. ASF22.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
CX NCBI_TaxID:63846;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ASF22.
RA Egan S., Weiner P., Kallifidas D., Wellington E.M.H.;
RT "Analysis of streptomycin biosynthetic gene clusters in streptomycetes
RT isolated from soil.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF064199; AAF75106.1;
DR HSSP: P00933; 2TYS
DR InterPro: IPR001926; B6_enzyme_beta.
DR Pfam: PF00291; PALP; 1.
FI NON_TER 1
FI NON_TER 66
SQ SEQUENCE 66 AA: 6908 MW: 97EB6A1C198CF618 CRC64;

Query Match 2.4%; Score 6; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 GLILVN 114
DB 50 GLILVN 55

RESULT 12
Q9N055
ID Q9N055 PRELIMINARY: PRT; 68 AA.
AC Q9N055
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DI 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
GN Unnamed protein product.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecidae; Macaca.
CX NCBI_TaxID:9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046036; BAB01618.1;
SQ SEQUENCE 68 AA: 7096 MW: 36F9D4A9FB28A029 CRC64;

Query Match 2.4%; Score 6; DB 6; Length 68;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 ALIKVV 148
DB 46 ALIKVV 51

RESULT 13
Q91QS3
ID Q91QS3 PRELIMINARY: PRT; 70 AA.
AC Q91QS3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DI 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE EBNA-1P protein (Fragment).
GN EBNA-1P.
OS Pongine herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
CX NCBI_TaxID:159604;
RN [1]
RP SEQUENCE FROM N.A.
RA McCann E.M.;
RT "A Genetic Analysis of Epstein-Barr Virus Coded Leader Protein EBNA-1P
RT As A Co-Activator of EBNA2 Function.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ311194; CAC40643.1;
DR InterPro: IPR005030; Herpes_LP.
DR Pfam: PF03363; Herpes_LP; 1.
FI NON_TER 1
FI NON_TER 70
SQ SEQUENCE 70 AA: 7580 MW: CD5E1C939805AF1F CRC64;

Query Match 2.4%; Score 6; DB 12; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

FT NON_TER 66
SQ SEQUENCE 66 AA: 6908 MW: 97EB6A1C198CF618 CRC64;

Query Match 2.4%; Score 6; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 GLILVN 114
DB 50 GLILVN 55

RESULT 12
Q9N055
ID Q9N055 PRELIMINARY: PRT; 68 AA.
AC Q9N055
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DI 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
GN Unnamed protein product.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecidae; Macaca.
CX NCBI_TaxID:9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046036; BAB01618.1;
SQ SEQUENCE 68 AA: 7096 MW: 36F9D4A9FB28A029 CRC64;

Query Match 2.4%; Score 6; DB 6; Length 68;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 ALIKVV 148
DB 46 ALIKVV 51

RESULT 13
Q91QS3
ID Q91QS3 PRELIMINARY: PRT; 70 AA.
AC Q91QS3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DI 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE EBNA-1P protein (Fragment).
GN EBNA-1P.
OS Pongine herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
CX NCBI_TaxID:159604;
RN [1]
RP SEQUENCE FROM N.A.
RA McCann E.M.;
RT "A Genetic Analysis of Epstein-Barr Virus Coded Leader Protein EBNA-1P
RT As A Co-Activator of EBNA2 Function.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ311194; CAC40643.1;
DR InterPro: IPR005030; Herpes_LP.
DR Pfam: PF03363; Herpes_LP; 1.
FI NON_TER 1
FI NON_TER 70
SQ SEQUENCE 70 AA: 7580 MW: CD5E1C939805AF1F CRC64;

Query Match 2.4%; Score 6; DB 12; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 220 GQCEPE 225
DB 41 GQCEPE 46

RESULT 14
Q9QV15
ID Q9QV15 PRELIMINARY; PRT; 8 AA.
AC Q9QV15;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DI 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TRENBLrel. 22, Last annotation update)
DE 16 kDa cell growth factor (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID:10118;
RN 11
RP SEQUENCE.
RX MEDLINE:92028975; PubMed:1656977;
RA Milner P.G.;
RT "Simian sarcoma virus transformation of normal rat kidney fibroblasts
RT is associated with markedly increased basic fibroblast growth factor
RT expression."
RL Biochem. Biophys. Res. Commun. 180:423-430(1991).
FT NON_TER 1
FT NON_TER 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 807 MW; H0787AAB07673AFA CRC64;

Query Match 2.0%; Score 5; DB 11; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 PEDGG 228
DB 4 PEDGG 8

RESULT 15
Q94785
ID Q94785 PRELIMINARY; PRT; 11 AA.
AC Q94785;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DI 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
DE Thrombopoietin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID:9606;
RN 11
RP SEQUENCE FROM N.A.
RA Sasaki Y., Takahashi T., Nakamura K., Okuno Y., Nakao K.;
RT "Production of thrombopoietic by Human Carcinomas and Its Novel mRNA
RT Isoforms."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AH014683; HAA34932.1;
FI NON_TER 1
FI NON_TER 1
SQ SEQUENCE 11 AA; 1203 MW; 5FE19F44B6C1A877 CRC64;

Query Match 2.0%; Score 5; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 SLIGT 85
DB 2 SLIGT 6

RESULT 16
P82444
ID P82444 PRELIMINARY; PRT; 16 AA.

```

```

P82444;
DT 01-JUN-2000 (TRENBLrel. 14, Created)
DI 01-JUN-2000 (TRENBLrel. 14, Last sequence update)
DE 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE 36 kDa cell wall protein (Fragment).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
CX NCBI_TaxID:4097;
RN 11
RP SEQUENCE.
RC STRAIN-CV. PETIT HAVANA;
RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA Wojtaszek P., Bolwell G.P.;
RT "Proteomic study of secondary cell wall proteins from transformed
RT tobacco culture."
RL Planta 200:0(2000).
CC -!- SUBCELLULAR LOCATION: CELL WALL.
CC -!- TISSUE SPECIFICITY: XYLEM.
KW Cell wall.
FT NON_TER 16
FT NON_TER 16
SQ SEQUENCE 16 AA; 1785 MW; FC98C09A51C9DF9C CRC64;

Query Match 2.0%; Score 5; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 GLQEG 132
DB 9 GLQEG 13

RESULT 17
Q905J6
ID Q905J6 PRELIMINARY; PRT; 19 AA.
AC Q905J6;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DI 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Tat protein (fragment).
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID:11676;
RN 11
RP SEQUENCE FROM N.A.
RC STRAN-96CG23;
RA Taniguchi Y., Takehsia J., Bikandou B., Mboudjeka I.,
RA N'Doundou-N'Kodia M.-Y., M'Pele P., Harada Y., Hayami M.,
RA Ichimura H., Parra H.-J.;
RT "Genetic Subtypes of HIV Type 1 Based on the vpu/unv Sequences in
RT Republic of Congo-Brazzaville."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF410440; AAL10226.1;
FI NON_TER 1
FI NON_TER 1
SQ SEQUENCE 19 AA; 2391 MW; 3E5E7ED9D0C038637 CRC64;

Query Match 2.0%; Score 5; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 RRPPE 160
DB 4 RRPPE 8

RESULT 18
Q9AUF0
ID Q9AUF0 PRELIMINARY; PRT; 21 AA.
AC Q9AUF0;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DI 01-JUN-2001 (TRENBLrel. 17, Last sequence update)

```

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE Stearoyl-ACP desaturase (Fragment).
 GN DE:AY BN-1.
 OS Brassica napus (Rape).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC Geraniales II; Brassicales; Brassicaceae; Brassica.
 CX NCBI_TaxID:3708;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN:CV. DRAKKAR;
 RA Fourmann M., Proger N., Brunel D.;
 RE "Amplified consensus gene markers: tools designing for a genetic map
 RE of Arabidopsis-known-function genes in Brassica";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EN5L; AF230688; AAK14945.1;
 FT NON-TER
 FI NON-TER 21 2;
 SQ SEQUENCE 21 AA; 2199 MW; 44C112B352HFB7F3 CRC64;
 Query Match 2.0%; Score 5; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 204 ASSSP 208
 Db 13 ASSSP 17
 [1]
 RESULT 19
 C9R5K3 PRELIMINARY; PRT; 23 AA.
 ID C9R5K3
 AC C9R5K3
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DI 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE GFP-GREEN flavoprotein.
 OS Proteobacteria; Alphaproteobacteria.
 CC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;
 CC Photobacterium.
 CX NCBI_TaxID:658;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=92081445; PubMed=1746316;
 RA Raibekas A.A.;
 RE "Green flavoprotein from P. leiognathi: purification, characterization
 RE and identification as the product of the lux G(N) gene";
 FI J. Bacteriol. Chemilumin. 6:169-176(1991).
 DR ASSP; P09342; INFP.
 DR InterPro; IPR002103; Bac. luciferase.
 DR Pfam; PF00296; Bac_luciferase; 1.
 SQ SEQUENCE 23 AA; 2806 MW; 5C8BCFAB1FBA368F CRC64;
 Query Match 2.0%; Score 5; DB 2; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 220 GQCEP 224
 Db 17 GQCEP 21
 [1]
 RESULT 20
 P82997 PRELIMINARY; PRT; 24 AA.
 ID P82997
 AC P82997
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DI 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Glutathione S-transferase (EC 2.5.1.18) (Fragment).
 OS Pseudomonas sp. (strain CF506).
 CC Bacteria; Proteobacteria.
 CX NCBI_TaxID 79676;

[1]
 RN SEQUENCE, FUNCTION, CATALYTIC ACTIVITY, SUBUNIT, AND SUBCELLULAR
 RP LOCATION.
 RX PubMed=11900268;
 RA Santos P.M., Mignogna G., Heipieper H.J., Zennaro E.;
 RE "Occurrence and properties of glutathione S-transferases in phenol-
 RE degrading Pseudomonas strains";
 RL Res. Microbiol. 153:89-98(2002).
 CC -|- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF
 CC EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
 CC -|- CATALYTIC ACTIVITY: RX - GLUTATHIONE - HX - R-S-GUTATHIONE.
 CC -|- SUBUNIT: MONOMER AND HOMODIMER.
 CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -|- SIMILARITY: BELONGS TO THE GST SUPERFAMILY.
 KW Transferase.
 FT NON-TER 24 24
 SQ SEQUENCE 24 AA; 2756 MW; 7C5BBA2B3C9C32BA CRC64;
 Query Match 2.0%; Score 5; DB 2; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 189 ALLEK 193
 Db 19 ALLEK 23
 [1]
 RESULT 21
 O49748 PRELIMINARY; PRT; 25 AA.
 ID O49748
 AC O49748
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DI 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Hypothetical 3.0 kDa protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20544085; PubMed=11094976;
 RA Irgentmann S.M.;
 RE "ERN1, a novel ethylene-regulated nuclear protein of Arabidopsis.";
 FI Plant Mol. Biol. 44:11-25(2000).
 DR EMBL; Y15066; CAA75350.1;
 KW Hypothetical protein.
 SQ SEQUENCE 25 AA; 2990 MW; 3AFAF166FA4B891E CRC64;
 Query Match 2.0%; Score 5; DB 10; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 40 IYIFD 44
 Db 11 IYIFD 15
 [1]
 RESULT 22
 Q9AUE5 PRELIMINARY; PRT; 25 AA.
 ID Q9AUE5
 AC Q9AUE5
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DI 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE Stearoyl-ACP desaturase (Fragment).
 GN DE:AY BN-1.
 OS Brassica campestris (Field mustard).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Brassica.
 CX NCBI_TaxID=3711;

```

RN SEQUENCE FROM N.A.
RP STRAIN-CV, R500;
RA Fournier M., Proger N., Brunel D.; Tools designing for a genetic map
RT of Arabidopsis known-function genes in Brassica.;
RL Submitted (FEB-2003) to the FMBI/GenBank/DOBC databases.
DR EMBL: AF240695; AAK14970.1;
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA: 2607 MW: AAN743B344B667F0 CRC64;

Query Match 2.0% Score 5; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 ASSSP 208
DB 17 ASSSP 21

RESULT 23
Q9T8S0
ID Q9T8S0 PRELIMINARY; PRT: 27 AA.
AC Q9T8S0;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE NADH dehydrogenase subunit 1 (Fragment).
GN NDI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RA LaFauci G., Kaczmarek W., Pagini M., Pullarkat R.K., Wisniewski K.E.,
RA Zhong N., Rubenstein R.;
RT Characterization of alternatively spliced transcripts of the Batten
RT disease CLN3 gene in human lymphoblastoid cell lines.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DOBC databases
DR EMBL: AF077973; AAK51495.1;
SQ SEQUENCE 27 AA: 2571 MW: EF8K915RA6731DE6 CRC64;

Query Match 2.0% Score 5; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 SDSFG 242
DB 12 SDSFG 16

RESULT 24
Q9T8S6
ID Q9T8S6 PRELIMINARY; PRT: 27 AA.
AC Q9T8S6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE NADH dehydrogenase subunit 1 (Fragment).
GN NDI.
OS Liolaemus capillatus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Tropidurinae; Liolaemus.
OX NCBI_TaxID:109431;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte J.A. II, Macey J.R., Espinoza R.E., Larson A.;
RT Phylogenetic relationships in the Iguanid lizard Genus Liolaemus;
RT Multiple origins of viviparous reproduction and evidence for recurring

```

```

RT Andean vicariance and dispersal.;
RC Biol. J. Linn. Soc. 69:75-102(2000).
DR EMBL: AF099234; AAF18806.1;
DR InterPro: IPR001694; Resp_NADH_dhl.
DR Pfam: PF00146; NADHdh; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA: 2969 MW: F876EC5D0744B1CD CRC64;

Query Match 2.0% Score 5; DB 8; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALLAT 6
DB 19 ALLAT 23

RESULT 25
Q9T8SC
ID Q9T8SC PRELIMINARY; PRT: 27 AA.
AC Q9T8SC;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE NADH dehydrogenase subunit 1 (Fragment).
GN NDI.
OS Liolaemus buergeri.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Tropidurinae; Liolaemus.
OX NCBI_TaxID:109400;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte J.A. II, Macey J.R., Espinoza R.E., Larson A.;
RT Phylogenetic relationships in the Iguanid lizard Genus Liolaemus;
RT Multiple origins of viviparous reproduction and evidence for recurring
RT Andean vicariance and dispersal.;
RL Biol. J. Linn. Soc. 69:75-102(2000).
DR EMBL: AF099236; AAF18812.1;
DR InterPro: IPR001694; Resp_NADH_dhl.
DR Pfam: PF00146; NADHdh; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA: 2999 MW: F861DC5D0744B1CD CRC64;

Query Match 2.0% Score 5; DB 8; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALLAT 6
DB 19 ALLAT 23

RESULT 26
Q9T8R8
ID Q9T8R8 PRELIMINARY; PRT: 27 AA.
AC Q9T8R8;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE NADH dehydrogenase subunit 1 (Fragment).
GN NDI.
OS Liolaemus cecil.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Tropidurinae; Liolaemus.
OX NCBI_TaxID:109442;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte J.A. II, Macey J.R., Espinoza R.E., Larson A.;

```


RT *Phylogenetic relationships in the iguanid lizard Genus *Liolepus*:
 RT Multiple origins of viviparous reproduction and evidence for recurring
 RT Andean vicariance and dispersal. ;
 RC Riosl. J. Linn. Soc. 69:75-102(2000).
 DK EMBL: AF098237; AAF18815.1; ;
 DR InterPro: IPR021694; Resp_NADH_dhl.
 DR Pfam: PF01446; NADBDH; ;
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 27 AA; 2999 MW; F861DC5D744B1CD CRC64;

Query Match 2.0%; Score 5; DB 8; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALLAT 6
 DB 19 ALLAT 23

RESULT 27
 Q9J574

ID Q9J574 PRELIMINARY; PRI: 27 AA.
 AC Q9J574;
 DT 01-MAY-2000 (TrEMBLrel. 15, Created)
 DT 01-MAY-2000 (TrEMBLrel. 15, Last sequence update)
 DI 01-MAY-2000 (TrEMBLrel. 15, Last annotation update)
 DE Leukotriene C4 synthase (Fragment).
 CS Mus sp.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20219428; PubMed-7766706;
 RA Coppelt-Strube M.;
 RT *Two step purification of human and murine leukotriene C4 synthase.*;
 RL Biochim. Biophys. Acta 1256:257-261(1995).
 SO SQUENCE 27 AA; 2936 MW; 9A4B5C7E143D23D1 CRC64;

Query Match 2.0%; Score 5; DB 11; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALLAT 6
 DB 6 ALLAT 10

RESULT 28
 Q9J574

ID Q9J574 PRELIMINARY; PRI: 27 AA.
 AC Q9J574;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
 DE (NS1)] (Fragment).
 CS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 CC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20219428; PubMed-10756048;
 RA Manzini A., Solforosi L., Debiaggi M., Zera F., Tanzi E., Romano L.,
 RA Zanetti A.R., Clementi M.;
 RT *Dominant role of host selective pressure in driving hepatitis C virus
 RT evolution in perinatal infection.*;
 RL J. Virol. 74:4327-4334(2000).
 DK EMBL: AF192415; AAF65626.1; ;
 DR InterPro: IPR002531; HCV_NS1;
 DR Pfam: PF01560; HCV_NS1; 1.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1
 SQ SEQUENCE 27 AA; 2775 MW; F9B1D42A0724BD64 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAOC 246
 DB 6 GGAOC 10

RESULT 29
 Q9J573

ID Q9J573 PRELIMINARY; PRI: 27 AA.
 AC Q9J573;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
 DE (NS1)] (Fragment).
 CS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 CC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20219428; PubMed-10756048;
 RA Manzini A., Solforosi L., Debiaggi M., Zera F., Tanzi E., Romano L.,
 RA Zanetti A.R., Clementi M.;
 RT *Dominant role of host selective pressure in driving hepatitis C virus
 RT evolution in perinatal infection.*;
 RL J. Virol. 74:4327-4334(2000).
 DK EMBL: AF192416; AAF65627.1; ;
 DR InterPro: IPR002531; HCV_NS1;
 DR Pfam: PF01560; HCV_NS1; 1.
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2722 MW; E5DDA42A0724BD64 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAOC 246
 DB 6 GGAOC 10

RESULT 30
 Q9J571

ID Q9J571 PRELIMINARY; PRI: 27 AA.
 AC Q9J571;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DI 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
 DE (NS1)] (Fragment).
 CS Hepatitis C virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 CC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20219428; PubMed-10756048;
 RA Manzini A., Solforosi L., Debiaggi M., Zera F., Tanzi E., Romano L.,
 RA Zanetti A.R., Clementi M.;
 RT *Dominant role of host selective pressure in driving hepatitis C virus

```

RT evolution in perinatal infection.*
RL J. Virol. 74:4327-4334(2000).
DR EMBL: AF192418; AAF65632.1; -.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 27 27
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2744 MW; F9B5642A0724BD64 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 31
Q9J5Z0
ID Q9J5Z0 PRELIMINARY; PRT; 27 AA.
AC Q9J5Z0
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20219428; PubMed=10756048;
RA Manzin A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
RA Zanetti A.R., Clementi M.;
RT "Dominant role of host selective pressure in driving hepatitis C virus
RT evolution in perinatal infection.*"
RL J. Virol. 74:4327-4334(2000).
DR EMBL: AF192418; AAF65632.1; -.
DR InterPro: IPR002531; HCV_NSL.
DR Pfam: PF01560; HCV_NSL.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 27 27
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2745 MW; F9B1C42B0724BD64 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 32
Q9J5Y9
ID Q9J5Y9 PRELIMINARY; PRT; 27 AA.
AC Q9J5Y9
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20219428; PubMed=10756048;
RA Manzin A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
RA Zanetti A.R., Clementi M.;
RT "Dominant role of host selective pressure in driving hepatitis C virus
RT evolution in perinatal infection.*"
RL J. Virol. 74:4327-4334(2000).
DR EMBL: AF192418; AAF65632.1; -.
DR InterPro: IPR002531; HCV_NSL.
DR Pfam: PF01560; HCV_NSL.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 27 27
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2745 MW; F9B1C42B0724BD64 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 33
Q9J5Y8
ID Q9J5Y8 PRELIMINARY; PRT; 27 AA.
AC Q9J5Y8
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20219428; PubMed=10756048;
RA Manzin A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
RA Zanetti A.R., Clementi M.;
RT "Dominant role of host selective pressure in driving hepatitis C virus
RT evolution in perinatal infection.*"
RL J. Virol. 74:4327-4334(2000).
DR EMBL: AF192418; AAF65632.1; -.
DR InterPro: IPR002531; HCV_NSL.
DR Pfam: PF01560; HCV_NSL.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 27 27
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2747 MW; 08D6D42A0724BD7C CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 34
Q9J5Y7
ID Q9J5Y7 PRELIMINARY; PRT; 27 AA.
AC Q9J5Y7
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

```

```

RA Manzin A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
RA Zanetti A.R., Clementi M.;
RT "Dominant role of host selective pressure in driving hepatitis C virus
RT evolution in perinatal infection.*"
RL J. Virol. 74:4327-4334(2000).
DR EMBL: AF192420; AAF65631.1; -.
DR InterPro: IPR002531; HCV_NSL.
DR Pfam: PF01560; HCV_NSL.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 27 27
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2775 MW; F9B1D42A0724BD64 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 33
Q9J5Y8
ID Q9J5Y8 PRELIMINARY; PRT; 27 AA.
AC Q9J5Y8
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20219428; PubMed=10756048;
RA Manzin A., Solforosi L., Debiaggi M., Zara F., Tanzi E., Romano L.,
RA Zanetti A.R., Clementi M.;
RT "Dominant role of host selective pressure in driving hepatitis C virus
RT evolution in perinatal infection.*"
RL J. Virol. 74:4327-4334(2000).
DR EMBL: AF192421; AAF65632.1; -.
DR InterPro: IPR002531; HCV_NSL.
DR Pfam: PF01560; HCV_NSL.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 27 27
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2747 MW; 08D6D42A0724BD7C CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 34
Q9J5Y7
ID Q9J5Y7 PRELIMINARY; PRT; 27 AA.
AC Q9J5Y7
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

```

```

OC Hepacivirus
OX NCBI_TaxID=11103;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE-20215428; PubMed-16756048;
RA Zanetti A., Sforzini L., Debiaggi M., Zara F., Tanzi F., Romero L.,
RA "Dominant role of host selective pressure in driving hepatitis C virus
RT evolution in perinatal infection.";
RI J. Virol. 74:4327-4334(2000).
DR EMBL: AF192422; AAF65633.1; .
DR InterPro: IPR002531; HCV_NSI; 1.
DR Pfam: PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2803 MW; E581E42A0724BD64 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 35
Q91K74 PRELIMINARY; PRT; 27 AA.
AC Q91K74;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein #2 (GP68) (GP70)
DE (NS1)] (Fragment).
GN POL.
GS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=LB1;
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
RA Rumeau J.L., Duffaut M., Durand D., Puel J.;
RC "Longitudinal analysis of Hepatitis C virus replication and liver
RT fibrosis progression in renal transplant recipients.";
RI Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF207328; AAF76569.1; .
DR InterPro: IPR002531; HCV_NSI; 1.
DR Pfam: PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2720 MW; A7E61601747D30C4 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 36
Q91K74 PRELIMINARY; PRT; 27 AA.
AC Q91K74;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP73)
DE (NS1)] (Fragment).
GN POL.
GS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=LB1;
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
RA Rumeau J.L., Duffaut M., Durand D., Puel J.;
RC "Longitudinal analysis of Hepatitis C virus replication and liver
RT fibrosis progression in renal transplant recipients.";
RI Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF207330; AAF76571.1; .
DR InterPro: IPR002531; HCV_NSI; 1.
DR Pfam: PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2720 MW; A7E61601747D30C4 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 37
Q91K73 PRELIMINARY; PRT; 27 AA.
AC Q91K73;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
GN POL.
GS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=LB1;
RA Izopet J., Rostaing L., Sandres K., Cisterne J.M., Pasquier C.,
RA Rumeau J.L., Duffaut M., Durand D., Puel J.;
RC "Longitudinal analysis of Hepatitis C virus replication and liver
RT fibrosis progression in renal transplant recipients.";
RI Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF207331; AAF76572.1; .
DR InterPro: IPR002531; HCV_NSI; 1.
DR Pfam: PF01560; HCV_NSI; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2720 MW; A7E61601747D30C4 CRC64;

Query Match 2.0%; Score 5; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GGAQG 246
DB 6 GGAQG 10

RESULT 38
Q91K74 PRELIMINARY; PRT; 27 AA.
AC Q91K74;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```

DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE putative cytochrome b (Fragment).
 GN PEIB2.
 OS Thiobacillus ferrooxidans.
 OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
 OX NCBI_TaxID=920;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SIMIAN-A:CC3020;
 RA Brusceila P., Levican G., Katouchniak J., Holmes D.S., Bonneloy V.:
 RT "A second operon encoding a bcl complex in Acidithiobacillus
 RT ferrooxidans.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/UDRJ databases.
 DR EMBL: AJ413196; CAC88363.1;
 FT NON_TIR 1
 SQ SEQUENCE 29 AA: 3445 MW: 7631FF992AC4401F CRC64:
 Query Match 2.0% Score 5; DB 2; Length 29;
 Best Local Similarity 100.0%; Pred. No. 1.6e-03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 31 PIRTP 35
 DB 16 PIRTP 20
 Search completed: April 10, 2003, 10:41:06
 Job time : 28.4706 secs

DT 01-DEC-2001 (TRENBLrel. 15, Created)
 DT 01-DEC-2001 (TRENBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Protease (Fragment).
 GN PRT.
 OS Simian retrovirus serogroup B.
 OC Viruses; Retrod viruses; Retroviridae.
 OX NCBI_TaxID=123562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-21011894; PubMed 1129642;
 RA Li R., Axthelm M.K., Machida C.A.:
 RT "Simian retrovirus serogroup B: partial gag-prt sequence and viral RNA
 RT distribution in an infected rhesus macaque.";
 RL Virus Genes 21:247-248(2000);
 DR EMBL: AF252389; AAF71356.1;
 FT NON_TIR 1
 SQ SEQUENCE 27 AA: 2767 MW: 83DF0E4D860224A0 CRC64:
 Query Match 2.0% Score 5; DB 15; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.5e-03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 228 GPAPG 232
 DB 15 GPAPG 19
 RESULT 39
 Q9N0N9 PRELIMINARY; PRT; 28 AA.
 AC Q9N0N9;
 DT 01-DEC-2000 (TRENBLrel. 15, Created)
 DT 01-DEC-2000 (TRENBLrel. 15, Last sequence update)
 DE Carcinoembryonic antigen-family cell adhesion molecule 1-1
 DE (Fragment).
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-21142399; PubMed 11245984;
 RA Zhou G.Q., Zhang Y., Hamadastron S.:
 RT "The carcinoembryonic antigen (CEA) gene family in non-human
 RT primates.";
 RL Gene 264:105-112(2001).
 DR EMBL: AF259568; AAF72663.1;
 FT NON_TIR 1
 SQ SEQUENCE 28 AA: 3067 MW: 6ERB8D78768CFB84 CRC64:
 Query Match 2.0% Score 5; DB 6; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.6e-03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 204 ASSSP 208
 DB 11 ASSSP 15
 RESULT 40
 Q93IG2 PRELIMINARY; PRT; 25 AA.
 ID Q93IG2
 AC Q93IG2

GenCore version 5.1.4 ps_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 10, 2003, 10:33:55 ; Search time 6.47059 Seconds
(without alignments)
1602.494 Million cell updates/sec

Title: US-09-930-026-1

Perfect score: 250

Sequence: 1 MALLA-SLPHG-MVKIFEDR.....PGEASGGSNSGGAGGLAFS 250

Scoring table: Oligo
Gapop 60.0 , Gapext 60.0

Searched: 1:2892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6856

Minimum DB seq length: 0

Maximum DB seq length: 73

Post processing: listing first 135 summaries

Database : SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	2.4	25	1	ENAT_MYCOA
2	5	2.0	15	1	HDSR_PIG
3	5	2.0	15	1	SPAH_HERAN
4	5	2.0	24	1	PCW5_PACCO
5	5	2.0	28	1	P771_HCMVT
6	5	2.0	36	1	Y260_BACDH
7	5	2.0	39	1	TAPD_ERWCA
8	5	2.0	42	1	Y282_THRPA
9	5	2.0	43	1	BAGL_HUMAN
10	5	2.0	47	1	THG1_MAIZE
11	5	2.0	53	1	ABAE_APINE
12	5	2.0	55	1	SARA_STRGC
13	5	2.0	58	1	V556_RPPE3
14	5	2.0	60	1	CX1_NAJWE
15	5	2.0	60	1	CX1_NAJMO
16	5	2.0	60	1	CX1_NAJPA
17	5	2.0	60	1	CX2_NAJOX
18	5	2.0	60	1	CX2_NAJMO
19	5	2.0	60	1	CX3_NAJMO
20	5	2.0	60	1	CX5T_NAJAT
21	5	2.0	69	1	CX0_NAJAT
22	5	2.0	61	1	CX1_HEMHA
23	5	2.0	61	1	CX2_HEMHA
24	5	2.0	61	1	CX2_NAJME
25	5	2.0	61	1	CX3_HEMHA
26	5	2.0	61	1	Y704_HAFIN
27	5	2.0	52	1	YM45_CAREL
28	5	2.0	63	1	Y442_ASISC
29	5	2.0	65	1	V07K_CLV
30	5	2.0	66	1	RPB1_CAEER
31	5	2.0	68	1	TXG5_CUFSA
32	4	1.6	8	1	AKHG_CRYHI
33	4	1.6	8	1	RST_MYCET

34	4	1.6	9	1	D1_NEPNO
35	4	1.6	9	1	NRE_HV128
36	4	1.6	10	1	CPA8_HUMAN
37	4	1.6	10	1	URE3_MORMO
38	4	1.6	11	1	QZOA_COMTE
39	4	1.6	12	1	CD11_LITXA
40	4	1.6	13	1	AU11_LITRA
41	4	1.6	13	1	AU12_LITRA
42	4	1.6	13	1	FIBB_HYLLA
43	4	1.6	13	1	CHAL_HUMAN
44	4	1.6	14	1	MAST_VESCR
45	4	1.6	15	1	CDN5_LITCE
46	4	1.6	15	1	CDN5_LITCE
47	4	1.6	15	1	FRE2_LITIN
48	4	1.6	15	1	SODM_STRGR
49	4	1.6	15	1	UC20_MAIZE
50	4	1.6	16	1	AU22_LITAU
51	4	1.6	16	1	AJ23_LITAU
52	4	1.6	16	1	AJ24_LITAU
53	4	1.6	16	1	AU25_LITRA
54	4	1.6	16	1	AJ26_LITRA
55	4	1.6	16	1	CT12_LITCI
56	4	1.6	16	1	CT13_LITCI
57	4	1.6	16	1	DH83_RHIDE
58	4	1.6	16	1	FIHA_EQUAS
59	4	1.6	16	1	PGT1_PELAC
60	4	1.6	17	1	AU31_LITRA
61	4	1.6	17	1	AJ32_LITRA
62	4	1.6	17	1	AJ33_LITRA
63	4	1.6	17	1	EPG_IHEAQ
64	4	1.6	17	1	LPW_AZOBK
65	4	1.6	17	1	PC24_BRANA
66	4	1.6	18	1	CT1C_LITCI
67	4	1.6	18	1	CT1D_LITCI
68	4	1.6	18	1	UC03_MAIZE
69	4	1.6	19	1	CH10_CLOPA
70	4	1.6	19	1	H3_NARPS
71	4	1.6	19	1	OXLA_OPHHA
72	4	1.6	19	1	UC31_MAIZE
73	4	1.6	20	1	ATP4_SPIOL
74	4	1.6	20	1	CAQS_RAT
75	4	1.6	20	1	CPA7_PAPSP
76	4	1.6	20	1	FIBB_ELEMA
77	4	1.6	20	1	MDH_KIHAR
78	4	1.6	21	1	NDK_CANAL
79	4	1.6	23	1	GP50_BPSP1
80	4	1.6	23	1	PS3_PSEPD
81	4	1.6	23	1	PS4_PSEPD
82	4	1.6	24	1	HPTA_RABIT
83	4	1.6	24	1	KAD_BACIL
84	4	1.6	24	1	RS19_PHYS2
85	4	1.6	25	1	ACP_ACICA
86	4	1.6	25	1	AMP3_MELGA
87	4	1.6	25	1	RS19_ACHLA
88	4	1.6	25	1	Y194_ARCFU
89	4	1.6	26	1	DHAB_AWME
90	4	1.6	26	1	LCIP_HUMAN
91	4	1.6	26	1	PORA_METTM
92	4	1.6	26	1	Y126_ARCFU
93	4	1.6	27	1	ACHA_MOUSE
94	4	1.6	27	1	ANF_ANQJA
95	4	1.6	27	1	DH81_BIFLO
96	4	1.6	27	1	L52_ADE07
97	4	1.6	27	1	RT10_BOVIN
98	4	1.6	28	1	PHYB_ASPEI
99	4	1.6	28	1	RS19_PHYS1
100	4	1.6	28	1	Y16P_HPT4
101	4	1.6	29	1	CU36_LCCHI
102	4	1.6	29	1	DMD_RAT
103	4	1.6	29	1	GLUC_LAMFL
104	4	1.6	29	1	H2B2_ECHES
105	4	1.6	29	1	PETN_ANASP
106	4	1.6	29	1	SCX1_ANDNA

p24816	neplrops no
p12481	human immun
p30094	homo sapien
p17339	morquellia
p80464	comamonas t
p56245	litoria ran
p82386	litoria ran
p82387	litoria ran
p14472	hylobates l
p40528	homo sapien
p01516	vespa crabr
p82077	litoria cae
p82078	litoria cae
p82022	litoria inf
p80733	streptomyce
p80626	zea mays (m
p82389	litoria aur
p82390	litoria aur
p82391	litoria aur
p82392	litoria ran
p82393	litoria ran
p81840	litoria cit
p81846	litoria cit
p80605	rhizobium l
p14449	equus asinu
p80563	peibacter
p82394	litoria ran
p82395	litoria ran
p82396	litoria ran
Q01697	thermus aqu
p50871	azospirillu
p81097	brassicica na
p81844	litoria cit
p81845	litoria cit
p80609	zea mays (m
p81338	clostridium
p80553	narctissus p
p81383	ophiophagus
p80637	zea mays (m
p80085	spinacia ol
p19633	rattus norv
p80055	papio sp. (
p14538	elephas max
p19578	kibdelospor
Q9ur66	candida alb
O48434	bacterioph
p83190	pseudis par
p83191	pseudis par
p13571	oryctolagus
p35140	bacillus li
O66096	phytoplasm
p80916	acinetobact
p80393	meleagris q
p29224	acholoplasm
O30045	archaeoglob
p86472	amycolatops
p34168	homo sapien
p80900	methanobact
O30111	archaeoglob
O70174	mus musculu
p18144	anguilla ja
p17615	bifidobacte
p35663	human adeno
p82670	bos taurus
p81440	aspergillus
O66093	phytoplasm
p39248	bacterioph
p17337	locusta mig
p11530	rattus norv
Q9prq5	lampetra fl
p13282	echinus esc
Q913p6	anabena sp
p56215	androctonus

107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135

4 1.6 29 1 1116_SPTOL
4 1.6 30 1 CXTA_CONF
4 1.6 30 1 PWB_BACNO
4 1.6 30 1 PSAM_OGOS1
4 1.6 30 1 R7A_SCHPO
4 1.6 30 1 RIPS_MOMCO
4 1.6 31 1 BCAM_PIG
4 1.6 31 1 COX4_NEUCR
4 1.6 31 1 MOE_STRAR
4 1.6 31 1 CTX2_CORVA
4 1.6 31 1 PETL_OGOS1
4 1.6 31 1 PSAM_EUGGR
4 1.6 31 1 Y3KO_BPCRHP
4 1.6 32 1 FRIH_ANAPL
4 1.6 32 1 GUR4_RAT
4 1.6 32 1 PUS5_DESRN
4 1.6 32 1 Y433_BORRO
4 1.6 32 1 BRZA_RANIS
4 1.6 33 1 PHNA_STPMU
4 1.6 33 1 PK1_DICDI
4 1.6 33 1 SC53_CANEA
4 1.6 34 1 CX35_CONGE
4 1.6 34 1 P15472_conus_geogr
4 1.6 35 1 C350_BACHA
4 1.6 35 1 P4666_bombyx_mori
4 1.6 35 1 CRCH_ANTIPE
4 1.6 35 1 GDRB_CLOPU
4 1.6 35 1 NSF_HVIE3
4 1.6 35 1 PSAL_CYAPA

P81834 spinacia ol
P58923 conus tulip
P17829 bacteroides
P49487 odontella s
P22667 schizosacch
P20655 momordica c
P19098 sus scrofa
P06809 neurospora
P82601 coriariocnem
P19982 streptomyce
P49524 odontella s
P31479 euglena gra
P19187 bacterioph
P80145 anas platyr
P33581 talius norv
P13064 desulfovibr
O51394 borrelia bu
P40837 rana escul
Q02419 streptococ
P34101 dictyosteli
P82008 canis fami
P31560 marchantia
P15472 conus geogr
P80091 bacillus ha
P4666 bombyx mori
P01509 antheraca p
P55793 clostridium
P05854 human immun
P48116 cyanophora

ALIGNMENTS

RESULT 1
ID DRAJ_MYCCA STANDARD; PRT; 25 AA.
AC 57-500;
DI 15-JUL-1998 (Rel. 36, Created)
DI 15-JUL-1998 (Rel. 36, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chaperone protein draj (fragment).
GN DRAJ.
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Entomoplasmatales.
OX NCBI TaxID=2935.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 25416;
RX MEDLINE=97148974; PubMed=8955799;
RA Falah M., Gupta R.S.;
RT Cloning of the draj (hsp70) gene region of Mycoplasma capricolum.*;
RC Int. J. Syst. Bacteriol. 47:38-45(1997).
CC 1- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC THE ATPASE ACTIVITY OF DRAJ (BY SIMILARITY).
CC 1- COFACTOR: BINDS TWO ZINC IONS PER MONOMER (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC 1- SIMILARITY: BELONGS TO THE DRAJ FAMILY.
CC 1- SIMILARITY: CONTAINS 1 J DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
EMBL: U51235; AAR09431.1;
DR InterPro: IPR001305; Draj_CXXCXXG.
DR InterPro: IPR001623; Draj_N.

DR Pfam: PF00226; Draj; 1.
DR PROSITE: PS00636; Draj_1; PARTIAL.
DR PROSITE: PS0076; Draj_2; PARTIAL.
DR PROSITE: PS00637; Draj_CXXCXXG; PARTIAL.
KW Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding.
FT DOMAIN 3 >25
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2915 MW; 430D30270BF0D0F1E CRC64;

Query Match 2.4%; Score 6; DB 1; Length 25;
Best Local Similarity 100.0%; Pred.No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 EGRIRQ 168
DB 18 EGRIRQ 23
| | | | |
| | | | |

RESULT 2
ID DGB_PIG STANDARD; PRI; 15 AA.
AC P20034;
DI 01-FEB-1991 (Rel. 17, Created)
DI 01-FEB-1991 (Rel. 17, Last sequence update)
DI 01-OCT-1996 (Rel. 34, Last annotation update)
DE Platelet-derived growth factor, B chain (PDGF B-chain) (PDGFR)
DE (Fragment).
GN PDGFR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI TaxID=9823;
RN [1]
RP SEQUENCE
RX MEDLINE=85126898; PubMed=6526008;
RA Stroobant P., Waterfield M.D.;
RT Purification and properties of porcine platelet-derived growth
RT factor.*;
RL EMBO J. 3:2963-2967(1984).
CC 1- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
CC 1- SUBUNIT: ANTI-PAALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
CC TRANSFORMATION PROCESSES.
CC 1- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
CC PDGF RECEPTOR.
CC 1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
DR PIR: A22789; A22789.
DR InterPro: IPR000072; PD_growth_factor.
DR PROSITE: PS00249; PDGF_1; PARTIAL.
DR PROSITE: PS0278; PDGF_2; PARTIAL.
KW Mitogen; Growth factor; Proto-oncogene; Platelet.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1411 MW; DCC39014450251C2 CRC64;

Query Match 2.0%; Score 5; DB 1; Length 15;
Best Local Similarity 100.0%; Pred.No. 2+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 PAAVE 215
DB 5 PAAVE 9
| | | | |
| | | | |

RESULT 3
ID SFAM_HELAN STANDARD; PRI; 18 AA.
AC P81098;
DI 15-JUL-1998 (Rel. 36, Created)
DI 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Seed fatty acyl-ester hydrolase (EC 3.1.1.1) (Fragment).
 OS Helianthus annuus (Common sunflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; eumetastadia 11; Asterales; Asteraceae; Asteroideae;
 CC Heliantheae; Helianthus.
 CX NCBI_TaxID=4232;
 RN [1]
 RP SEQUENCE.
 RC STRAIN CV. RUSTICA / VAR. EUROFOR; TISSUE=Seed;
 RA Beisson M., Gardies A.-M., Teissere M., Ferte N., Noat G.;
 RT "An esterase co-purified in post-germinated sunflower seeds is
 RT related to a new family of lipolytic enzymes.";
 RL plant physiol. Biochem. 35:761-765(1997).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=95210327; PubMed=7596323;
 RA Teissere M., Borel M., Cailliot R., Nari J., Gardies A.-M., Noat G.;
 RT "Purification and characterization of a fatty acyl-ester hydrolase
 RT from post-germinated sunflower seeds.";
 RL Biochim. Biophys. Acta 1255:105-112(1995).
 CC !- FUNCTION: IMPlicated in the BREAKDOWN OF OIL BODY-STORED LIPIDS
 CC DURING POST-GERMINATION.
 CC !- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O -> an alcohol + a
 CC carboxylic anion.
 CC !- ISSUES SPECIFICITY: SEED.
 CC !- DEVELOPMENTAL STAGE: POST-GERMINATION.
 CC !- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.
 DR InterPro: IP003108; Lipase_GDSL.
 DR PROSITE: PS01098; LIPASE_GDSL; SER; PARTIAL.
 KW Hydrolase; Lipid degradation; Glycoprotein.
 FT AC SITE 13 13 BY SIMILARITY.
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2657 MW; 3E9A80EB6548B62 CRC64;

 Query Match 2.0%; Score 5; DB 1; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2.4e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 72 LYDNG 76 2.0%; Score 5; DB 1; Length 18;
 DB 14 LYDNG 19

 RESULT 4
 PCW5_PACGO STANDARD; PRT; 24 AA.
 AC PR2427;
 DT 16-OCT-2001 (Rel. 43, Created)
 D; 16-OCT-2001 (Rel. 43, Last sequence update)
 D; 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Penicillin W5
 OS Pachycondyla goeldii (Ponerine ant).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 OC Aculeata; Formicidae; Ponerinae; Pachycondyla.
 CX NCBI_TaxID=118888;
 RN [1]
 RP SEQUENCE, AND FUNCTION.
 RC TISSUE=Venom;
 RX MEDLINE=21264562; PubMed=11279030;
 RA Orivo J., Rodeker V., Le Caer J.-P., Krier F., Revol-Junelles A.-M.,
 RA Longuen A., Chafotte A., Dejean A., Rossier J.;
 RT "Ponerins, the new antibacterial and insecticidal peptides from the
 RT venom of the ant Pachycondyla goeldii.";
 RL J. Biol. Chem. 276:17823-17829(2001).
 CC !- FUNCTION: BROAD SPECTRUM OF ACTIVITY AGAINST BOTH GRAM-POSITIVE
 CC AND GRAM-NEGATIVE BACTERIA AND S.CERVISIAE. HAS .NSPECIFICIDAL.
 CC AND HEMOLYTIC ACTIVITIES.
 CC !- MASS SPECTROMETRY: MW=2559.67; MHMOD=MHMOD.
 KW Antibiotic; Insect immunity; Fungicide; Hemolysis.
 SQ SEQUENCE 24 AA; 2600 MW; DDEBFACAA7U71D7E CRC64;

Query Match 2.0%; Score 5; DB 1; Length 24;
 Best Local Similarity 100.0%; Pred. No. 3.1e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 26 ALIKG 30 2.0%; Score 5; DB 1; Length 24;
 DB 4 ALIKG 8

 RESULT 5
 PP71_HCMVT STANDARD; PRT; 28 AA.
 AC P24429;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-MAR-1992 (Rel. 21, Last annotation update)
 DE 71 kDa upper matrix phosphoprotein (pp71) (Fragment).
 GN UL82.
 OS Human cytomegalovirus (strain Towne).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Betaherpesvirinae; Cytomegalovirus.
 CX NCBI_TaxID=10363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220654; PubMed=1850902;
 RA Pande H., Campo K., Tanamachi H., Zaia J.A.;
 RT "Human cytomegalovirus strain Towne pp65 gene: nucleotide sequence
 RT and expression in Escherichia coli.";
 RL Virology 182:220-228(1991).
 CC !- SIMILARITY: BELONGS TO THE UL82 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CX
 DR EMBL: M67443; AAA45995.1;
 KW Phosphorylation; Matrix protein.
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 2554 MW; 3C96A69C384425BD CRC64;

 Query Match 2.0%; Score 5; DB 1; Length 28;
 Best Local Similarity 100.0%; Pred. No. 3.5e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 204 ASSSP 208 2.0%; Score 5; DB 1; Length 28;
 DB 4 ASSSP 8

 RESULT 6
 Y260_BACHD STANDARD; PRT; 36 AA.
 AC Q9KG53;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein BH0260.
 GN BH0260.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;

RT Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 RT halodurans and genomic sequence comparison with *Bacillus subtilis*.
 RT Nucleic Acids Res. 28:4317-4331 (2000).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AP001597; BAR049791.2.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 36 AA: 4307 MW: 62503504F98CEE27 CRC64;

Query Match: 2.0%; Score 5; DB 1; Length 36;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 111 ILVNE 115
 DB 3 ILVNE 7

RESULT 7
 TRPO_EHWC
 AC P12320; STANDARD: PRI; 39 AA.
 DT 01-OCT-1989 (Rel. 12, Created)
 DI 01-OCT-1989 (Rel. 12, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Anthranilate phosphoribosyltransferase (EC 2.4.2.18) (Fragment).
 OS *Erwinia carotovora*.
 CC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae;
 CC Pectobacterium.
 CX NCBI_TaxID=554;
 RN 11
 RP SEQUENCE.

RC STRAIN:RP9;
 RX MEDLINE:78066891; PubMed=338606;
 RA Lardon M., Mills S.E., Rowe J., Yanofsky C.;
 RT Purification and properties of a third form of anthranilate-5-
 R phosphoribosylpyrophosphate phosphoribosyltransferase from the
 RT Enterobacteriaceae.
 RJ J. Biol. Chem. 253:409-412 (1978).

CC -1- CATALYTIC ACTIVITY: Anthranilate + phosphoribosylpyrophosphate -
 CC N 5'-phosphoribosyl-anthranilate + diphosphate.
 CC -1- PATHWAY: Tryptophan biosynthesis: diphosphate.
 CC -1- SUPPLEMENT: HOMO:RUR.
 CC -1- SIMILARITY: BELONGS TO THE ANTHRANILATE PHOSPHORIBOSYLTRANSFERASE
 CC FAMILY.
 CC PR: AC3003; A05003.
 DR EMBL: F02885; Glycos_transf_3.
 DR Pfam: F02885; Glycos_transf_3; 1.
 KW Tryptophan biosynthesis; Transferase; Glycosyltransferase.
 FT NCL_TaxID=554;
 FT NCL_TaxID=554;
 SQ SEQUENCE 39 AA: 4459 MW: 62503504F98CEE27 CRC64;

Query Match: 2.0%; Score 5; DB 1; Length 39;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 QSMTP 153
 DB 24 QSMTP 28

RESULT 8
 Y28L_TRPRA
 ID Y28L-TRPRA STANDARD: PRI; 42 AA.
 AC O83305;

DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein TP0281.
 GN TP0281.
 OS *Treponema pallidum*.
 CC Bacteria: Spirochaetales: Spirochaetaceae; Treponema.
 CX NCBI_TaxID=160;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN:Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstein G.M., White G., Sutton G.G.,
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khaki H., Richardson D., Howell J.K., Chidambaram M., Utterback J.,
 RA McDonald L., Artach P., Howman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT Complete genome sequence of *Treponema pallidum*, the syphilis
 RT spirochete.*
 RL Science 281:375-388 (1998).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

EMBL: AE001209; AAC65278.1;
 TIGR: TP0281;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 42 AA: 4462 MW: F9FAD73F78BCD7F CRC64;

Query Match: 2.0%; Score 5; DB 1; Length 42;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ELSDS 219
 DB 30 ELSDS 34

RESULT 9
 BAGE_HUMAN
 ID BAGE_HUMAN STANDARD: PRI; 43 AA.
 AC Q13072;
 DI 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DE 16-OCT-2002 (Rel. 40, Last annotation update)
 DE B melanoma antigen (Antigen M22-HA).
 GN BAGE.
 OS Homo sapiens (Human).
 CC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
 CC Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo.
 CX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=Melanoma;
 RX MEDLINE=95202592; PubMed=7895173;
 RA Roel P., Wildmann C., Sensi M.L., Brasseur K., Renaud J.-C.,
 RA Coulic P., Boon T., van der Bruggen P.;
 RT BAGE: a new gene encoding an antigen recognized on human melanomas
 RT by cytolytic T lymphocytes.*
 RL Immunity 2:167-175 (1995).

CC -1- FUNCTION: NOT KNOWN. ANTIGEN RECOGNIZED ON A MELANOMA BY
 CC AUTOLOGOUS CYTOLYTIC T LYMPHOCYTES.
 CC -1- TISSUE SPECIFICITY: NOT EXPRESSED IN NORMAL TISSUES, EXCEPT IN
 CC TESTIS. EXPRESSED WITH SIGNIFICANT PROPORTION IN MELANOMAS, BUT
 CC ALSO IN TUMORS OF VARIOUS HISTOLOGICAL ORIGINS, SUCH AS BLADDER
 CC CARCINOMAS, HEAD AND NECK SQUAMOUS CELL CARCINOMAS, LUNG AND


```

CC BREAST CARCINOMAS, NOT EXPRESSED IN RENAL, COLORECTAL AND
CC PROSTATIC CARCINOMAS, LEUKEMIAS AND LYMPHOMAS. MORE FREQUENTLY
CC EXPRESSED IN METASTATIC MELANOMAS THAN IN PRIMARY MELANOMAS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U14183; AAC50123.1; -
DR GenBank: HGNC:942; BAGE.
DR MIM: 605167; -
DR InterPro: IPR001230; Proxyl_site.
KW Antigen.
SV SEQUENCE: 43 AA; 4910 MW; 36F3B8CF4312F1HB CRC64;
Query Match 2.0%; Score 5; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 223 EPDGG 227
DE 1 11
DE 32 EPDGG 36
RESULT 10
THGL_MAIZE
ID THGL_MAIZE STANDARD: PRT: 47 AA.
AC P87009;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-zeathionin 1.
OS Zea mays (Maize).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACG clade;
CC Panicoideae; Andropogoneae; Zea.
GX NCBI_TaxID:4577;
RN [1]
RP SEQUENCE.
RC ISSUES-Seed;
RA Castro M.S., Fortes W., Morhy L., Bloch C. J.F.;
RT "Complete amino acid sequences of two gamma-thionins from maize (Zea
RT mays L.) seeds".
RL Protein Rept. Lett. 3:267-274(1996).
CC -!- FUNCTION: THIONINS ARE SMALL PLANT PROTEINS WHICH ARE TOXIC
CC TO ANIMAL CELLS.
CC -!- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
DR HSSP: P20158; IGPS.
DR MaizeDB: L30775; -
DR InterPro: IPR002116; Gamma-thionin.
DR Pfam: PF00364; Gamma-thionin; 2.
DR ProDom: PD022594; Gamma-thionin; 1.
DR SMART: SM00505; KuoC1; 1.
DR ProSITE: PS00340; GAMMA THIONIN; 2.
KW Plant defense; plant toxin.
FT DISULFID 3 47 BY SIMILARITY.
FT DISULFID 14 34 BY SIMILARITY.
FT DISULFID 20 42 BY SIMILARITY.
FT DISULFID 24 43 BY SIMILARITY.
SV SEQUENCE 47 AA; 5199 MW; 0F3A74A58C3B8DFE CRC64;
Query Match 2.0%; Score 5; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 5.6e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 129 LOEY 133
DB 25 LOEY 25

```

```

RESULT 11
ABAE_APIME
ID ABAE_APIME STANDARD: PRT: 53 AA.
AC P15450;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update);
DE Abaecin precursor.
DE Abaecin (Honeybee).
OS Apis mellifera (Honeybee).
CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
CC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
CC Aculeata; Apoidea; Apidae; Apis.
GX NCBI_TaxID:7460;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:95050655; PubMed-7961803;
RA Casteels-Josson K., Zhang W., Capaci T., Casteels P., Tempst P.;
RT "Acute transcriptional response of the honeybee peptide-antibiotics
RT gene repertoire and required post-translational conversion of the
RT precursor structures".
RL J. Biol. Chem. 269:28569-28575(1994).
RN [2]
RP SEQUENCE OF 20-53.
RX ISSUP-Hemolymph;
RX MEDLINE:90126848; PubMed-2298215;
RA Casteels P., Anpe C., Riviere L., van Damme J., Elisone C., Jacobs F.,
RA Tempst P.;
RT "Isolation and characterization of abaecin, a major antibacterial
RT response peptide in the honeybee (Apis mellifera)".
RL Eur. J. Biochem. 187:381-386(1990).
CC -!- FUNCTION: THIS PEPTIDE HAS BACTERICIDAL ACTIVITY.
CC -!- SIMILARITY: PARTIAL TO APIAECINS AND DIPTERICINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U15954; AAA67442.1; -
DR PIR: S08152; S08152.
KW Insect immunity; Antibiotic; Hemolymph; Signal.
FT SIGNAL 1 19
FT CHAIN 20 53 ABAECIN.
FT VARIANT 52 52 G->S.
SV SEQUENCE 53 AA; 5903 MW; 0E1F0CC57797FBC CRC64;
Query Match 2.0%; Score 5; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 6.3e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ALLAT 6
DB 9 ALLAT 13
RESULT 12
SARA_STRGC
ID SARA_STRGC STANDARD: PRT: 55 AA.
AC P31306;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 76 kDa cell surface lipoprotein precursor (fragment).
GX NCBI_TaxID:29390;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:95050655; PubMed-7961803;
RA Casteels-Josson K., Zhang W., Capaci T., Casteels P., Tempst P.;
RT "Acute transcriptional response of the honeybee peptide-antibiotics
RT gene repertoire and required post-translational conversion of the
RT precursor structures".
RL J. Biol. Chem. 269:28569-28575(1994).
RN [2]
RP SEQUENCE OF 20-53.
RX ISSUP-Hemolymph;
RX MEDLINE:90126848; PubMed-2298215;
RA Casteels P., Anpe C., Riviere L., van Damme J., Elisone C., Jacobs F.,
RA Tempst P.;
RT "Isolation and characterization of abaecin, a major antibacterial
RT response peptide in the honeybee (Apis mellifera)".
RL Eur. J. Biochem. 187:381-386(1990).
CC -!- FUNCTION: THIS PEPTIDE HAS BACTERICIDAL ACTIVITY.
CC -!- SIMILARITY: PARTIAL TO APIAECINS AND DIPTERICINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U15954; AAA67442.1; -
DR PIR: S08152; S08152.
KW Insect immunity; Antibiotic; Hemolymph; Signal.
FT SIGNAL 1 19
FT CHAIN 20 53 ABAECIN.
FT VARIANT 52 52 G->S.
SV SEQUENCE 53 AA; 5903 MW; 0E1F0CC57797FBC CRC64;
Query Match 2.0%; Score 5; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 6.3e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ALLAT 6
DB 9 ALLAT 13

```

```

RN SEQUENCE FROM N.A.
RX MEDLINE: 92175970; PubMed-1339436;
RT Carlsberg H.F.;
RT Adherence, coaggregation, and hydrophobicity of Streptococcus
RT gordonii associated with expression of cell surface lipoproteins.*;
RL Infect. Immun. 60:1225-1228(1992).
CC -! FUNCTION: MAY BE INVOLVED IN THE EXPRESSION OF CELL SURFACE
CC PROPERTIES IMPORTANT FOR COLONIZATION OF THE HUMAN ORAL CAVITY.
CC IT MAY ALSO BE INVOLVED IN UPTAKE PROCESSES.
CC -! SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor.
CC -! SIMILARITY: WITH THE N-TERMINAL REGION OF S.PNEUMONIAE AMIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S85398; AAB21606.1;
DR PIR: A43896; A43896.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KW Membrane; lipoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 >55 76 KDA CELL SURFACE LIPOPROTEIN.
FT LIPID 23 23 N-ACYL DIGLYCERIDE.
FT NON-TER 55 55
FT SEQUENCE 55 AA; 5685 MW; C14536037536F1 CRC64;
SQ SEQUENCE 55 AA; 5685 MW; C14536037536F1 CRC64;

Query Match 2.0%; Score 5; DR 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 6.5e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALLAT 6
DB 13 ALLAT 17

RESULT 13
VGS8_BDPF3 STANDARD; PRT; 58 AA.
AC P03629;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE 6.4 kDa protein (DRF 58).
OS Bacteriophage Pf3.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID:10872;
RN SEQUENCE FROM N.A.
RX STRAIN New-York, and Nijmegen;
RX MEDLINE-85293231; PubMed-3928901;
RA Luiten R.G.M., Puterman D.G., Schoenmakers J.C.G., Konings R.N.H.,
RA Day L.A.;
RT "Nucleotide sequence of the genome of Pf3, an IncP-1 plasmid-specific
RT filamentous bacteriophage of Pseudomonas aeruginosa.*";
RL J. Virol. 56:268-276(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M19377; AAA88386.1;
DR EMBL: M11912; AAA88377.1;
DR PIR: A04237; 25BP83.
SQ SEQUENCE 58 AA; 6446 MW; 92C7C97502E12946 CRC64;

```

```

Query Match 2.0%; Score 5; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 6.8e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SLEPG 11
DB 25 SLEPG 29

RESULT 14
CXI_NAJME STANDARD; PRT; 60 AA.
ID CXI_NAJME
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 1 (Cytotoxin V-II-1) (Toxin V(II)1).
OS Naja melanoleuca (Forest cobra) (Black-lipped cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scieroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID:8644;
RN SEQUENCE.
RC TISSUE-Venom;
RA Carlsson F.H.H., Joubert F.J.;
RT "Snake venom toxins. The isolation and purification of three
RT cytotoxin homologues from the venom of the forest cobra (Naja
RT melanoleuca) and the complete amino acid sequence of toxin V(II)1.*";
RC Biochim. Biophys. Acta 336:453-469(1974).
CC -! SURCELLULAR LOCATION: Secreted.
CC -! TISSUE SPECIFICITY: Expressed by the venom gland.
CC -! MISCELLANEOUS: L5(50) IS 1.37 MG/KG BY INTRAVENOUS INJECTION.
CC -! SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR HSP: A01714; HSNJW.
DR HSP: P01449; 2CDX.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; Toxin; 1.
DR PRINTS: PR00282; CYTOTOXIN.
DR ProDom: PD000206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family.
FT DISULFID 3 21 BY SIMILARITY.
FT DISULFID 14 38 BY SIMILARITY.
FT DISULFID 42 53 BY SIMILARITY.
FT DISULFID 54 59 BY SIMILARITY.
FT VARIANT 1 1 L->I (IN EQUAL AMOUNT).
FT SEQUENCE 60 AA; 6682 MW; 6AD08FAH80BC44EA CRC64;
SQ SEQUENCE 60 AA; 6682 MW; 6AD08FAH80BC44EA CRC64;

Query Match 2.0%; Score 5; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 7e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 KSSLL 101
DB 44 KSSLL 48

RESULT 15
CXI_NAJMO STANDARD; PRT; 60 AA.
ID CXI_NAJMO
AC P01467;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 1 (Cardiotoxin XIIB) (Cytotoxin V-II-1) (CTX IIB).
OS Naja mossambica (Mozambique cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scieroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID:8644;

```

```

RN [1]
RP SEQUENCE.
RC TISSUE-Venom;
RA LOW A.I.;
RT "Snake venom toxins. The amino acid sequences of three cytotoxin
  homologues from Naja mossaambica venom.";
RL Biochim. Biophys. Acta 336:481-495(1974).
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE-88029481; PubMed-282421;
RA Oling G., Steinmetz W.E., Bougis P.E., Rochat H., Wuehrlich K.;
RT "Sequence-specific 1H-NMR assignments and determination of the
  secondary structure in aqueous solution of the cytotoxins CTXIIa
  and CTXIIb from Naja mossaambica venom.";
RL Eur. J. Biochem. 168:693-697(1987).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE-93275339; PubMed-8504828;
RA O'Connell C.F., Bougis P.E., Wuehrlich K.;
RT "Determination of the nuclear-magnetic-resonance solution structure
  of cardiac toxin CTX IIP from Naja mossaambica venom.";
RL Eur. J. Biochem. 213:891-900(1993).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: LD(50) IS 0.93 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR PIR: A01728; H3NJ1B.
DR PIR: S02517; S02517.
DR PDB: 2CCX; 31-JAN-94.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; toxin; 1.
DR PRINTS: PR00282; CYTOTOXIN.
DR PRODOM: PR00206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family; 3D-structure.
FT DISULFID 3 21
FT DISULFID 14 38
FT DISULFID 42 53
FT DISULFID 54 59
FT STRAND 2 3
FT TURN 9 10
FT TURN 12 13
FT STRAND 16 17
FT TURN 21 26
FT STRAND 30 31
FT TURN 35 37
FT STRAND 49 54
FT TURN 57 58
SQ SEQUENCE 60 AA; 6826 MW; 68350H77686491C CMC64;
Query Match 2.0%; Score 5; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 KSSL 101
DB 44 KSSL 48
IIII
IIII
RESULT 10
CX2_NAJMO STANDARD; PRT: 60 AA.
AC P01468;
DI 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 1 (Cardiotoxin gamma)
OS Naja pallida (Red spitting cobra)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Elapinae; Naja.
NCBI_TaxID=8658;

```

```

RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RC TISSUE-Venom;
RX MEDLINE-75205552; PubMed-1146181;
RA Fryklund L., Eaker D.;
RT "The complete covalent structure of a cardiotoxin from the venom of
  Naja nigricollis (African black-necked spitting cobra).";
RL Biochemistry 14:2865-2871(1975).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.55 ANGSTROMS).
RX MEDLINE-94254084; PubMed-8196041;
RA Bilwes A., Rees B., Moras D., Menez R., Menez A.;
RT "X-ray structure at 1.55 A of toxin gamma, a cardiotoxin from Naja
  nigricollis venom. Crystal packing reveals a model for insertion into
  membranes.";
RL J. Mol. Biol. 239:122-136(1994).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
CC -1- CAUTION: THE VENOM OF THIS SNAKE WAS ORIGINALLY THOUGHT TO BE THAT
  OF N. NIGRICOLLIS WHILE IT IS REALLY FROM N. PALLIDA.
DR PIR: A37578; H3NJ1B.
DR PDB: 1IGX; 30-APR-94.
DR PDB: 1CXN; 20-DEC-94.
DR PDB: 1CXO; 20-DEC-94.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; toxin; 1.
DR PRINTS: PR00282; CYTOTOXIN.
DR PRODOM: PR00206; Snake_toxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family; 3D-structure.
FT DISULFID 3 21
FT DISULFID 14 38
FT DISULFID 42 53
FT DISULFID 54 59
FT STRAND 2 4
FT TURN 8 9
FT STRAND 11 13
FT TURN 16 17
FT STRAND 20 26
FT TURN 27 28
FT STRAND 29 29
FT TURN 30 31
FT STRAND 35 39
FT STRAND 49 54
FT TURN 57 58
SQ SEQUENCE 60 AA; 6827 MW; 68B50B996B6491C CMC64;
Query Match 2.0%; Score 5; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 KSSL 101
DB 44 KSSL 48
IIII
IIII
RESULT 17
CX2_NAJMO STANDARD; PRT: 60 AA.
AC P01469;
DI 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 2 (Cardiotoxin XIa) (Cytotoxin V-II-2)
OS Naja mossaambica (Mozambique cobra)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Elapidae; Elapinae; Naja.
NCBI_TaxID=8644;
RN [1]
RP SEQUENCE.

```

RC TISSUE Venom:
RA Zoon A.1.;
RT "Snake venom toxins. The amino acid sequences of three cytotoxin
RL homologues from Naja mossambica mossambica venom.*;
RN Biochim. Biophys. Acta 336:481-495(1974).
RN [2]
RX STRUCTURE BY NMR.
RX MEDLINE:88029481; PubMed-282421;
RA Oetting G., Stelmach W.E., Bougis P.E., Kochat H., Kuehrich K.;
RT "Sequence-specific 1H NMR assignments and determination of the
RC secondary structure in aqueous solution of the cardiotoxins CTX1a
R2 and CTX1b from Naja mossambica mossambica.*;
R2 Eur. J. Biochem. 168:609-620(1987).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: LD(50) IS 1.1 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR PIR: AC1723; H3NJ2M.
DR PIR: S02518; S02518.
DR HSP: P01468; ITGX.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; Toxin_1.
DR PRINTS: PK0282; CYTOTOXIN.
DR ProDom: PD00206; Snake_toxin_1.
DR PROSITE: PS00272; SNAKE_TOXIN_1.
KW Venom: Cytotoxin; Cardiotoxin; Multigene family.
FT DISULFID 1 21 BY SIMILARITY.
FT DISULFID 14 38 BY SIMILARITY.
FT DISULFID 42 53 BY SIMILARITY.
FT DISULFID 54 59 BY SIMILARITY.
SQ SEQUENCE 60 AA; 6600 MW; 42940C1996B024A CRC64;
Query Match 2.0%; Score 5; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 KSSL 101
DB 44 KSSL 48
RESULT 18
CX2_NAJOX
ID CX2_NAJOX STANDARD; PRI; 60 AA.
AC P0144;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 2.
CS Naja oxiana (Central Asian cobra).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
CC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=8657;
RN [2]
RP SEQUENCE.
RC "SSU-Venom:
RX MEDLINE:75C8179; PubMed-4435217;
RA Grishin E.V., Sukhikh A.P., Adamovich I.B., Ovchinnikov Y.A.,
RA Yukelson L.Y.;
RT "The isolation and sequence determination of a cytotoxin from the
RT venom of the Middle-Asian cobra Naja oxiana.*;
RL FEBS Lett. 48:179-183(1974).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR PIR: A01799; H3NJ2R.
DR HSP: P01444; 2CR.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; Toxin_1.
DR PRINTS: PK0282; CYTOTOXIN.

DR ProDom: PD00206; Snake_toxin_1.
KW Venom: Cytotoxin; Cardiotoxin; Multigene family.
FT DISULFID 3 21 BY SIMILARITY.
FT DISULFID 14 38 BY SIMILARITY.
FT DISULFID 42 53 BY SIMILARITY.
FT DISULFID 54 59 BY SIMILARITY.
SQ SEQUENCE 60 AA; 6636 MW; 3512FDB5ED2C5F7 CRC64;
Query Match 2.0%; Score 5; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 KSSL 101
DB 44 KSSL 48
RESULT 19
CX3_NAJMO
ID CX3_NAJMO STANDARD; PRI; 60 AA.
AC P01470;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 3 (Cytotoxin V-II-3).
CS Naja mossambica (Mozambique cobra).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
CC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=8644;
RN [1]
RP SEQUENCE.
RC TISSUE-Venom:
RA Louw A.1.;
RT "Snake venom toxins. The amino acid sequences of three cytotoxin
RL homologues from Naja mossambica mossambica venom.*;
RL Biochim. Biophys. Acta 336:481-495(1974).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: LD(50) IS 1.82 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR PIR: A01730; H3NJ3M.
DR HSP: P01467; 2CCX.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam: PF00087; Toxin_1.
DR PRINTS: PK0282; CYTOTOXIN.
DR ProDom: PD00206; Snake_toxin_1.
DR PROSITE: PS00272; SNAKE_TOXIN_1.
KW Venom: Cytotoxin; Cardiotoxin; Multigene family.
FT DISULFID 3 21 BY SIMILARITY.
FT DISULFID 14 38 BY SIMILARITY.
FT DISULFID 42 53 BY SIMILARITY.
FT DISULFID 54 59 BY SIMILARITY.
SQ SEQUENCE 60 AA; 6894 MW; 01E1F647BDD9FAE CRC64;
Query Match 2.0%; Score 5; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 KSSL 101
DB 44 KSSL 48
RESULT 20
CX57_NAJAT
ID CX57_NAJAT STANDARD; PRI; 60 AA.
AC P07525;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cytotoxin 5 (Cardiotoxin analog V) (CTX-5) (CTX V) (Cardiotoxin 1)
 DE (Cytotoxin D1) (Membrane toxin D1).
 OS Naja atra (Chinese cobra).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 CC Elapidae; Elapinae; Naja.
 OX NCBI_TaxID=8656;
 RN [1]
 RP SEQUENCE
 RC TISSUE=Venom;
 RA Wu Y.-L., Du Y.-C.;
 RT "The complete amino acid sequence of cytotoxin D-1";
 RL Acta Biochim. Biophys. Sin. 15:310-315(1984).
 RN [2]
 RP STRUCTURE BY NMR.
 RC TISSUE=Venom;
 RA MEDLINE=20252380; PubMed=10784406;
 RA Jayaraman G., Kumar T.K.S., Tsui C.-C., Srisaillam S., Chou S.-H.,
 Hu C.-Z., Yu C.;
 RT "Elucidation of the solution structure of cardiotoxin analogue V from
 the Taiwan cobra (Naja atra) -- identification of structural
 features important for the lethal action of snake venom
 cardiotoxins";
 RC Protein Sci. 9:637-646(2000).
 CC -1- FUNCTION: THIS TOXIN IS CARDIOTOXIC AND CYTOTOXIC TO YOSHIDA
 SARCOMA CELLS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
 DR PIR: JCQ001; R3N5SF.
 DR PDB: 1CHV; 30-MAR-00.
 DR InterPro: IPR003572; Cytotoxin.
 DR Pfam: PF00087; Snake_toxin.
 DR PRINTS: PR00282; CYTOTOXIN.
 DR ProDom: PD000206; Snake_toxin; 1.
 DR PROSITE: PS00272; Snake_toxin; 1.
 KW Venom; Cytotoxin; Cardiotoxin; Calcium channel inhibitor;
 K 3D-structure; Multigene family.
 FT DISULFID 3 21
 FT DISULFID 14 38
 FT DISULFID 42 53
 FT DISULFID 54 59
 SQ SEQUENCE 60 AA: 6810 MW: 51305F8E5CB78F1C CRC64;

Query Match 2.0%; Score 5; DB 1; Length 60;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 KSSLL 101
 DB 44 KSSLL 48

RESULT 22
 CX6_NAJAT STANDARD; PRT; 60 AA.
 ID CX6_NAJAT
 DE 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytotoxin 6 (Cardiotoxin 6) (CTX6).
 OS Naja atra (Chinese cobra).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 CC Elapidae; Elapinae; Naja.
 OX NCBI_TaxID=8656;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RA MEDLINE=34251167; PubMed=8193587;
 RA Hung C.-C., Wu S.-H., Chion S.-H.;
 RT "Sequence characterization of cardiotoxins from Taiwan cobra:

RT isolation of a new isoform";
 RL Biochem. Mol. Biol. Int. 31:1031-1040(1993).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
 DR HSSP: P01444; 2CPI.
 DR InterPro: IPR003572; Cytotoxin.
 DR InterPro: IPR003571; Snake_toxin.
 DR Pfam: PF00087; toxin; 1.
 DR PRINTS: PR00282; CYTOTOXIN.
 DR ProDom: PD000206; Snake_toxin; 1.
 DR PROSITE: PS00272; SNAKE_TOXIN; 1.
 KW Venom; Cytotoxin; Cardiotoxin; Multigene family.
 FT DISULFID 3 21 BY SIMILARITY.
 FT DISULFID 14 38 BY SIMILARITY.
 FT DISULFID 42 53 BY SIMILARITY.
 FT DISULFID 54 59 BY SIMILARITY.
 SQ SEQUENCE 60 AA: 6689 MW: 1E754BE220B2AEEC CRC64;

Query Match 2.0%; Score 5; DB 1; Length 60;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 KSSLL 101
 DB 44 KSSLL 48

RESULT 22
 CX1_HEMHA STANDARD; PRT; 61 AA.
 ID CX1_HEMHA
 DE 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytotoxin 1 (Hemolytic protein 12B).
 OS Hemachatus haemachatus (Ringhals) (Sepedon haemachatus).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 CC Elapidae; Elapinae; Hemachatus.
 OX NCBI_TaxID=8626;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RA MEDLINE=73140212; PubMed=4734897;
 RA Fyklund L., Baker D.;
 RT "Complete amino acid sequence of a nonneurotoxic hemolytic protein
 from the venom of Haemachatus haemachatus (African ringhals cobra).";
 RL Biochemistry 12:661-667(1973).
 CC -1- FUNCTION: THIS PROTEIN Lyses RED BLOOD CELLS AND HAS CARDIOTOXIC
 AND HYPOTENSIVE ACTIVITIES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
 DR PIR: A01731; H3R1LL.
 DR HSSP: P01444; 2CPI.

DR InterPro: IPR003572; Cytotoxin.
 DR InterPro: IPR003571; Snake_toxin.
 DR Pfam: PF00087; toxin; 1.
 DR PRINTS: PR00282; CYTOTOXIN.
 DR ProDom: PD000206; Snake_toxin; 1.
 DR PROSITE: PS00272; SNAKE_TOXIN; 1.
 KW Venom; Cytotoxin; Cardiotoxin; Multigene family.
 FT DISULFID 3 22 BY SIMILARITY.
 FT DISULFID 15 39 BY SIMILARITY.
 FT DISULFID 43 54 BY SIMILARITY.
 FT DISULFID 55 60 BY SIMILARITY.
 SQ SEQUENCE 61 AA: 6844 MW: EDF501522BB76087 CRC64;

Query Match 2.0%; Score 5; DB 1; Length 61;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 97 KSSIL 101
DB 45 KSSIL 49

RESULT 24
CX2_NAME CX2_HEMHA STANDARD: PRT: 61 AA.
AC P24777:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 2 (Toxin 12A).
CS Hemachatus haemachatus (Ringhals) (Sapedon haemachatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Hemachatus.
OX NCBI_TaxID=8626;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=77162036; PubMed 404150;
RA Joubert F.J.;
RT *Snake venom toxins. The amino-acid sequences of three toxins (9B, 11 and 12A) from Hemachatus haemachatus (Ringhals) venom.*;
RL Eur. J. Biochem. 74:387-395(1977).
CC -1- FUNCTION: THIS PROTEIN LYSSES RED BLOOD CELLS AND HAS CARDIOTOXIC AND HYPOTENSIVE ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: LD(50) IS 3.8 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR HSSP; P14554; IXXI.
DR InterPro: IPR003572; Cytotoxin.
DR Pfam: PF00087; toxin; 1.
DR PRINTS; PR00282; CYTOTOXIN.
DR PRODOM; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; Snake_toxin; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 15 39 BY SIMILARITY.
FT DISULFID 43 54 BY SIMILARITY.
FT DISULFID 55 60 BY SIMILARITY.
SQ SEQUENCE 61 AA; 6803 MW; A40DFE3390014B CRC64;

Query Match 2.3%; Score 5; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 7,1e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 KSSIL 101
DB 45 KSSIL 49

RESULT 24
CX2_NAME CX2_HEMHA STANDARD: PRT: 61 AA.
AC P24777:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 3 (Toxin 11/11A).
CS Hemachatus haemachatus (Ringhals) (Sapedon haemachatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Hemachatus.
OX NCBI_TaxID=8626;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=77162036; PubMed 404150;
RA Joubert F.J.;
RT *Snake venom toxins. The amino-acid sequences of three toxins (9B, 11 and 12A) from Hemachatus haemachatus (Ringhals) venom.*;
RL Eur. J. Biochem. 74:387-395(1977).
CC -1- FUNCTION: THIS PROTEIN LYSSES RED BLOOD CELLS AND HAS CARDIOTOXIC AND HYPOTENSIVE ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: LD(50) IS 2.5 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR HSSP; P01444; 2CST.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR PRINTS; PR00282; CYTOTOXIN.
DR PRODOM; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 15 39 BY SIMILARITY.

```

```

RT *Snake venom toxins. The primary structures of two novel cytotoxin
RT homologues from the venom of forest cobra (Naja melanoleuca)*;
RL Biochem. Biophys. Res. Commun. 59:269-276(1974).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: THESE TWO SEQUENCES HAVE GLY AND LEU AT POSITIONS
CC 25 AND 27, RESPECTIVELY, INSTEAD OF 2 MET THAT ARE FOUND IN MOST
CC OTHER CYTOTOXINS AND CARDIOTOXINS. THE LOW TOXICITY OF THESE 2
CC PROTEINS MAY POSSIBLY BE CORRELATED WITH THESE CHANGES.
CC -1- MISCELLANEOUS: LD(50) IS 6.25 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR PIR; A01734; H3N2W.
DR HSSP; P14554; IXXI.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR PRINTS; PR00282; CYTOTOXIN.
DR PRODOM; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 15 39 BY SIMILARITY.
FT DISULFID 43 54 BY SIMILARITY.
FT DISULFID 55 60 BY SIMILARITY.
SQ SEQUENCE 61 AA; 6850 MW; C80D099A6D6A530B CRC64;

Query Match 2.0%; Score 5; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 7,1e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 KSSIL 101
DB 45 KSSIL 49

RESULT 25
CX3_HEMHA STANDARD: PRT: 61 AA.
AC P24777:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytotoxin 3 (Toxin 11/11A).
CS Hemachatus haemachatus (Ringhals) (Sapedon haemachatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Hemachatus.
OX NCBI_TaxID=8626;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=77162036; PubMed 404150;
RA Joubert F.J.;
RT *Snake venom toxins. The amino-acid sequences of three toxins (9B, 11 and 12A) from Hemachatus haemachatus (Ringhals) venom.*;
RL Eur. J. Biochem. 74:387-395(1977).
CC -1- FUNCTION: THIS PROTEIN LYSSES RED BLOOD CELLS AND HAS CARDIOTOXIC AND HYPOTENSIVE ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- MISCELLANEOUS: LD(50) IS 2.5 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR HSSP; P01444; 2CST.
DR InterPro: IPR003572; Cytotoxin.
DR InterPro: IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR PRINTS; PR00282; CYTOTOXIN.
DR PRODOM; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Cardiotoxin; Multigene family.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 15 39 BY SIMILARITY.

```

FT DISULFID 43 54 BY SIMILARITY.
 FT DISULFID 55 60 BY SIMILARITY.
 FT VARIANT 10 10 F -> Y (IN TOXIN 11A).
 SQ SEQUENCE 61 AA: 6793 MW: 3F73CC72588225EC CRC64;

Query Match 2.0%; Score 5; DB 1; Length 61;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 KSSL 101
 Db 45 KSSL 49

RESULT 26
 Y704_HAEIN STANDARD; PRI: 61 AA.
 AC 04434C;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein H10734.
 GN H10704

OS Haemophilus influenzae;
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
 OC Haemophilus.

OX NCBI_TaxID=727;
 RN [1]

RP SEQUENCE FROM N.A.

RC S-RAIN-ad / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischman R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.K., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKeown K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-F., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs L., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.I., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gelfand C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT *Whole-genome random sequencing and assembly of Haemophilus influenzae
 WT Rd.;

KL Science 259:456-512(1995).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: U32753; AAC22367.1; -;
 DR IIGK: H10704; -;

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 61 AA: 6302 MW: DDD5512ED25773AF CRC64;

Query Match 2.0%; Score 5; DB 1; Length 61;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 196 ALPNC 200

Db 27 ALPNC 31

RESULT 27

YMA5_CAEEL STANDARD; PRI: 62 AA.

ID YMA5_CAEEL

AC P34521;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical 7.0 kDa protein KILH3.5 in chromosome III.
 GN KILH3.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Rhabditinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=94150718; PubMed=7905398;

RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonnenhammer E., Stader R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Wellstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.;

KL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: Z22180; CAA80177.1; -;

DR PIR: S40758; S40758.

DR WormPep: KILH3.5; CR00265.

KW Hypothetical protein.

SQ SEQUENCE 62 AA: 7047 MW: 75FC9PC791608AB5 CRC64;

Query Match 2.0%; Score 5; DB 1; Length 62;

Best Local Similarity 100.0%; Pred. No. 7.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 ALLEK 193

Db 7 ALLEK 11

RESULT 28

CXH2_ASPSC

ID CXH2_ASPSC STANDARD; PRI: 63 AA.

AC P19003;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cytotoxin homolog S3C2

OS Aspidelaps scutatus (Shield-nose snake).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosaurs; Squamata; Scleroglossa; Serpentes; Colubroidea;

OC Elapidae; Acanthophiinae; Aspidelaps.

OX NCBI_TaxID=8607;

RN [1]

RP SEQUENCE.

RC TISSUE-Venom;

RX MEDLINE=88185648; PubMed=3356299;

RA Joubert F.J.;

RT *Snake venom toxins -- II. The primary structures of cytotoxin

RT homologues S3C2 and S4C8 from Aspidelaps scutatus (shield or

RT shield-nose snake) venom.;

RL Int. J. Biochem. 20:337-345(1988).

CC -I- SUBCELLULAR LOCATION: Secreted.

CC -I- TISSUE SPECIFICITY: Expressed by the venom gland.

CC -I- MISCELLANEOUS: LD(50) IS 6.6 MG/KG BY INTRAVENOUS INJECTION.

```

CC -!- SIMILARITY: BELONGS TO THE SNAKE TOXIN FAMILY.
DR PIR: JSC298; JSC298.
DR HSSP: P01444; 2CKT.
DR InterPro: IPR003572; Cytotoxin.
DR Pfam: PF00087; Toxin; 1.
DR PRNTS: PR00082; CYTOTOXIN.
DR ProDom: PD00026; SnakeToxin; 1.
DR PROSITE: PS00272; SNAKE_TOXIN; 1.
KW Venom; Cytotoxin; Multigene family.
FT DISULFID 3 22 BY SIMILARITY.
FT DISULFID 25 40 BY SIMILARITY.
FT DISULFID 44 55 BY SIMILARITY.
FT DISULFID 56 61 BY SIMILARITY.
SQ SEQUENCE 63 AA; 7173 MW; 18733F42E9D1F270 CRC64;

Query Match 2.0%; Score 5; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 KSSLL LCL
DB 45 KSSLL 50

RESULT 29
V07K_CIV STANDARD; PRT: 65 AA.
AC V288V8;
DT 01 DEC-1992 (Rel. 24, Created)
DT 01 DEC-1992 (Rel. 24, Last sequence update)
DT 01 DEC 1994 (Rel. 30, Last annotation update)
DE 7 kDa protein (ORF 4).
CS Carnation latent virus (CLV).
CC Viruses; ssRNA positive-strand viruses, no DNA stage; Carlaviruses.
OX NCBI_TaxID: 2164;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91355918; PubMed: 2103483;
RA Foster G.D., Meenan B.M., Mills P.R.;
RT "Nucleotide sequence of the 7K gene of carnation latent virus.*";
RL Plant Mol. Biol. 15:937-939(1990).
CC -!- SIMILARITY: TO OTHER 7 kDa PROTEINS (ORF4) FROM POTEXVIRUSES
AND CARLAVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC FMBL: X55331; CAA39030.1; -
DR PIR: S12405; S12405.
DR InterPro: IPR003411; Coat_7kDa.
DR Pfam: PF02455; 7kD_coat; 1.
KW Transmembrane.
SQ SEQUENCE 65 AA; 6843 MW; 57EFAC7226FD2A06 CRC64;

Query Match 2.0%; Score 5; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 GLILV 1:3
DB 5 GLILV 9

RESULT 30
RPBL_CAE38 STANDARD; PRT: 66 AA.
ID RPBL_CAE38
AC P35074;

```

```

DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase II largest subunit (PC 2.7.7.6)
DE (fragment).
GN AMA-1.
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID: 6238;
RN [1]
RP SEQUENCE FROM N.A.
RX Bird D.M., Wilson M.A., Kaioshian I.;
RT Submitted (XX-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate -
CC [RNA](N).
CC -!- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM NINE TO
CC FOURTEEN DIFFERENT POLYPEPTIDES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC THE PHOSPHORYLATION ACTIVATES POL2.
CC -!- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND 5.8S rRNA GENES.
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC FMBL: L23763; AAA27891.1; -
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
FT NON_TER 66
SQ SEQUENCE 66 AA; 7237 MW; DCEE4449DF1C9E0B CRC64;

Query Match 2.0%; Score 5; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GPAPG 232
DB 55 GPAPG 59

RESULT 31
TXC9_CUPSA STANDARD; PRT: 68 AA.
ID TXC9_CUPSA
AC P58604;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toxin CSTX-9.
OS Cupiennius salei (Wandering spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Cupiennius.
OX NCBI_TaxID: 6928;
RN [1]
RP SEQUENCE, DISULFIDE BONDS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RT TISSUE-Venom.
RX MEDLINE: 21544892; PubMed: 21693532;
RA Schaller J., Kacemper U., Schuerch S., Kuhn-Mentwig L., Haeblerli S.,
RA Mentwig W.;
RT *CSTX-9, a toxic peptide from the spider Cupiennius salei; amino acid
RT sequence, disulphide bridge pattern and comparison with other spider

```



```

R# Toxins containing the cysteine knot structure.*;
RL Cell. Mol. Life Sci. 55:1538-1545(2001)
CC -!- FUNCTION: THIS TOXIN CAUSES PARALYSIS IN DROSOPHILA WITH AN LD50
CC VALUE OF 3.12 PMOL/MG.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- MASS SPECTROMETRY: MW:7529.75; MW_ERR:0.32; METHOD:Electrospray.
KW Toxin; Neurotoxin.
FT DISULFID 5 21
FT DISULFID 13 30
FT DISULFID 20 48
FT DISULFID 32 46
SQ SEQUENCE 68 AA; 7539 MW; 5B3526DCA359F9F4 CRC64;
Query Match: 2.3%; Score 5; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 98 SSLLQ 102
DB 50 SSLLQ 54
RESULT 32
AKHG GRVBI
ID AKHG GRVBI STANDARD; PRT; 8 AA.
AC 214086;
DT 01-FEB-1994 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE Adipokinetic hormone G (AKH-G) (KO II).
CS Glycyls binaculatus (Two-spotted cricket), and
OS Romalea microptera (Lubber grasshopper).
CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
CC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera;
CC Gryllidae; Gryllinae; Gryllus.
CC NCBI_TaxID=6959, 7007;
CC 11;
CC SEQUENCE.
CC SPECIES-G.bimaculatus: TISSUE-Corpora cardiaca;
RX MEDLINE=8H106553; PubMed=3425616;
RA Gaele G., Winchert K.L., Jr.;
RT "Primary sequence analysis by fast atom bombardment mass spectrometry
RT of a peptide with adipokinetic activity from the corpora cardiaca of
RL the cricket Gryllus bimaculatus.*;
RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
RX 11;
RX SEQUENCE.
RX SPECIES-R.microptera: TISSUE-Corpora cardiaca;
RX MEDLINE=89145002; PubMed=3226948;
RA Gaele G., Hilbich G., Beyreuther K., Riechert K.L., Jr.;
RT "Sequence analyses of two neuropeptides of the AKH/RPCH-family from
RT the lubber grasshopper, Romalea microptera.*;
RL Peptides 9:681-688(1989).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / URTH / RPCH FAMILY.
DR INTERPRO: IPR002047; AKH.
DR PIR: A28004; A28004.
DR PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;
Query Match: 1.6%; Score 4; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 FSTG 173
1111

```

```

DB 4 FSTG 7
RESULT 33
RS7_MYCIT
ID RS7_MYCIT STANDARD; PRT; 8 AA.
AC P33564;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 30S ribosomal protein S7 (Fragment).
GN RPS3.
OS Mycobacterium intracellulare.
CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC NCBI_TaxID=1767;
RN 11;
RP SEQUENCE FROM N.A.
RX MEDLINE=93197130; PubMed=8451173;
RA Nair J., Rouse D.A., Morris S.L.;
RT "Nucleotide sequence analysis of the ribosomal S12 gene of
RT Mycobacterium intracellulare.*;
RL Nucleic Acids Res. 21:1039-1039(1993).
CC -!- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF
CC 16S RIBOSOMAL RNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: 108171; AAA25376.1;
DR PIR: S35538; S35538.
DR INTERPRO: IPR000235; Ribosomal_S7.
DR PROSITE: PS00052; RIBOSOMAL_S7; PARTIAL.
KW Ribosomal protein; rRNA-binding.
FT INIT_MET 0 0 BY SIMILARITY.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 850 MW; 63276DC768732417 CRC64;
Query Match: 1.6%; Score 4; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 228 GPAP 231
1111
DB 4 GPAP 7
RESULT 34
DL_NEPNO
ID DL_NEPNO STANDARD; PRT; 9 AA.
AC P24816;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Gastrin/cholecystokinin-like peptide D1.
OS Nephrops norvegicus (Norway lobster).
CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
CC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
CC Asiacidea; Nephropoidea; Nephropidae; Nephrops.
CC NCBI_TaxID=6829;
RN 11;
RP SEQUENCE.
RX MEDLINE=92082847; PubMed=1747388;
RA Favrel P., Kegel G., Sedlmeier B., Keller R., van Wormhoudt A.;
RT "Structure and biological activity of crustacean gastrointestinal
RT peptides identified with antibodies to gastrin/cholecystokinin.*;

```

RL Biochimie 73:1233-1239(1991);
 CC -/- FUNCTION: MAY CONTROL DIGESTION PROCESSES IN CRUSTACEANS.
 CC -/- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR PIR: A48398; A48398.
 KW Hormone.
 SQ SEQUENCE 9 AA: 1038 MW: 636C79CA36D8787B CRC64:

Query Match 1.6% Score 4; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 SHG 243
 III
 DB 1 SEQ 4

RESULT 35

NEF_HV178
 ID NEF_HV178 STANDARD; PRT: 9 AA.
 AC P12481;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Negative factor (F-protein) (27 kDa protein) (Fragment).
 GN NEF.
 OS Human immunodeficiency virus type 1 (Z-84 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID:11681;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE:88281278; PubMed:3395517;
 RA YOUNG J., JOSEPHS S.F., REITZ M.S. JR., ZAGURY D., WONG-STAL F.,
 RA GALLO R.C.;
 RT "Nucleotide sequence analysis of the env gene of a new Zairian
 isolate of HIV-1";
 RI AIDS Res. Hum. Retroviruses 4:165-173(1988).
 CC -/- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
 activities. It seems to down-regulate the CD4(T4) antigen.
 CC -/- MISCELLANEOUS: THE Z-84 ISOLATE WAS TAKEN FROM A 54 YEAR-OLD
 ZAIRIAN MALE.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

DR IMBL: J03653; AAA44687.1;
 DR HIV: J03653; NEFSJY1.
 KW AIDS; Myristate: GTP-binding.
 FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA: 967 MW: 319CH325A373387B CRC64:

Query Match 1.6% Score 4; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e-05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 SKSS 99
 III
 DB 6 SKSS 9

RESULT 36

UPA8_HUMAN
 ID UPA8_HUMAN STANDARD; PRT: 10 AA.
 AC P30094;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-SEP-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of plasma (Spot 34) (Fragment).

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN 11

RP SEQUENCE

RC TISSUF-Plasma;
 RX MEDLINE:93029337; PubMed:1459097;
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
 RA Hochstrasser D.F.;
 RI "Plasma protein map: an update by microsequencing";
 RI Electrophoresis 13:707-714(1992).

CC -/- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 7.2. ITS MW IS: 16 kDa.

DR SWISS-2DPAGE; P30094; HUMAN.

FT NON_TER 1 1
 FT VARIANT 4 4 S -> H.

FT NON_TER 10 10 /FTID=VAR_000003.

FT SEQUENCE 10 AA: 977 MW: 2EA6E0C77AE325B8 CRC64:

Query Match 1.6% Score 4; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.4e-03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 PAVP 55
 III
 DB 7 PAVP 10

RESULT 37

URE3_MORMO
 ID URE3_MORMO STANDARD; PRT: 10 AA.
 AC P17339;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Urease gamma subunit (EC 3.5.1.5) (6 kDa subunit) (Urea
 DE amidohydrolase) (Fragment).
 GN UREA.
 OS Morganella morganii (Proteus morganii).
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Morganella.
 OX NCBI_TaxID:582;
 RN 11
 RP SEQUENCE

RX MEDLINE:90264298; PubMed:2345135;
 RA Hu L.-T., Nicholson E.B., Jones R.D., Lynch M.J., Mobley H.L.I.;
 RT "Morganella morganii urease: purification, characterization, and
 RT isolation of gene sequences";
 RL J. Bacteriol. 172:3073-3080(1990).
 CC -/- CATALYTIC ACTIVITY: Urea + H(2)O -> CO(2) + 2 NH(3).
 CC -/- SUBUNIT: (ALPHA, BETA, GAMMA)(3) (BY SIMILARITY).
 CC -/- SIMILARITY: BELONGS TO THE UREASE GAMMA SUBUNIT FAMILY.
 DR PIR: C35389; C35389.
 KW Hydrolase.
 FT NON_TER 10 10

FT SEQUENCE 10 AA: 1171 MW: 4B313BCB07771A7 CRC64:

Query Match 1.6% Score 4; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.4e-03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 PPEV 161
 III
 DB 5 PPEV 8

RESULT 38

Q20A_COMTE
 ID Q20A_COMTE STANDARD; PRT: 11 AA.
 AC P80464;

DT 01-NOV-1995 (Rel. 32, Created)
 DI 01-NOV-1995 (Rel. 32, Last sequence update)
 DI 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinolone 2-oxido-reductase, alpha chain (EC 1.3.99.17) (Fragment).
 OS Comamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria: Proteobacteria: beta subdivision: Comamonadaceae; Comamonas.
 CX NCBI_TaxID=285;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=63;
 RX MEDLINE=96035889; PubMed=7556204;
 RA Schach S., Tshisuaka B., Fetzner S., Lingens F.;
 R1 "Quinolone 2-oxido-reductase and 2-oxo-1,2-dihydroquinoline 5,6-
 RT dioxigenase from Comamonas testosteroni 63. The first two enzymes in
 RI quinoline and 3-methyl-quinoline degradation.";
 RL Eur. J. Biochem. 232:536-544(1995).
 CC [1]
 CC [1] FUNCTION: CONVERTS (3-METHYL-)QUINOLINE TO (3-METHYL-)2-OXO-
 CC 1,2-DIHYDROQUINOLINE.
 CC [1] CATALYTIC ACTIVITY: Quinolone + acceptor + H(2)O -> isoquinoline-
 CC 1(2H)-one + reduced acceptor.
 CC [1] COFACTOR: Fe, MOLYBDENUM AND IRON-SULFUR.
 CC [1] PATHWAY: FIRST STEP IN THE DEGRADATION OF QUINOLINE AND
 CC (3-METHYL-)QUINOLINE.
 CC [1] SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND
 CC TWO GAMMA CHAINS (PROBABLE).
 KW Oxido-reductase; Flavoprotein; FAD; Molybdenum.
 KW NON-TER 1;
 SQ SEQUENCE 21 AA: 1213 MW: 869094322B1DC2CA CRC64;

 Query Match 1.6%; Score 4; DB 1; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 213 VAPL 216
 DB 5 VAPL 8

 RESULT 39
 CD_LITRA
 ID CD_LITRA STANDARD; PRT; 12 AA.
 AC P56245; P91253;
 DT 15-JUL-1994 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Caeridin 1.1/1.2/1.3.
 OS Litoria xanthomera (Orange-thighed frog).
 OS Litoria splendida (Magnificent tree frog).
 OS Litoria gleni (Common tree frog), and
 OS Litoria chloris (Blue-thighed frog).
 OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 OC Amphibia: Batrachia: Anura: Neobatrachia: Bufonidae; Hylidae;
 OC Pelodyadinae: Litoria.
 CX NCBI_TaxID=79697, 30345, 39405, 86064;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC SPECIES=L.xanthomera;
 RX MEDLINE=97374030; PubMed=5230483;
 RA Steinbörner S.T., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.,
 RA Ramsay S.;
 R1 "New caerin antibacterial peptides from the skin glands of the
 RT Australian tree frog Litoria xanthomera.";
 RL J. Pept. Sci. 3:181-185(1997).
 CC [2]
 CC [2] SPECIES=L.splendida; TISSUE=Parotoid gland;
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
 R1 "Peptides from Australian frogs. Structures of the caerins and
 RI caeridin 1 from Litoria splendida.";
 RL J. Chem. Soc. Perkin Trans. 1:3173-3176(1992).
 CC [3]
 CC [3] FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST B.CEREUS, L.LACTIS,
 CC L.INNOCUA AND S.UBERIS. PROBABLY ACTS BY DISRUPTING MEMBRANE
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC SPECIES=L.gilleni; TISSUE=Parotoid gland;

RA Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;
 R1 "Peptides from Australian frogs. The structures of the caerins and
 RI caeridins from Litoria gilleni.";
 RL J. Chem. Res. 139:937-961(1993).
 RN [4]
 RP SEQUENCE, AND MASS SPECTROMETRY (CAERIDINS 1.1; 1.2 AND 1.3).
 RC SPECIES=L.gilleni;
 RA Waugh R.J., Steinbörner S.T., Bowie J.H., Wallace J.C., Tyler M.J.,
 RA Hu P., Gross M.L.;
 R1 "Two isomeric alpha and beta aspartyl dodecapeptides and their
 RT cyclic amino succinyl analogue from the Australian tree frog
 RI Litoria gilleni.";
 RL Aust. J. Chem. 46:1981-1987(1995).
 RN [5]
 RP SEQUENCE.
 RC SPECIES=L.chloris; TISSUE=Skin;
 RX MEDLINE=98175802; PubMed=9515047;
 RA Steinbörner S.T., Currie G.J., Bowie J.H., Wallace J.C., Tyler M.J.;
 R1 "New antibiotic caerin 1 peptides from the skin secretion of the
 RI Australian tree frog Litoria chloris. Comparison of the activities of
 RL the caerin 1 peptides from the genus Litoria.";
 RL J. Pept. Res. 51:121-126(1998).
 CC [1] FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTIHISTAMIC ACTIVITY.
 CC [1] TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS IN
 CC L.XANTHOMERA OR SPECIFICALLY BY THE SKIN PAROTOID AND/OR ROSTRAL
 CC GLANDS IN L.SPLENDIDA AND L.GILLIENI.
 CC [1] PTM: ISOMERIZATION ALPHA-BETA OF THE ASP-4 RESIDUE IN CAERIDIN
 CC 1.2; A CYCLIC SUCCINIMIDE MAY BE FORMED BETWEEN ASP-4 AND GLY-5
 CC RESIDUES IN CAERIDIN 1.3.
 CC [1] MASS SPECTROMETRY: MW=1140; METHOD=FAH.
 CC [1] MISCELLANEOUS: THE RESULTS FOR MASS SPECTROMETRY OF CAERIDIN 1.1
 CC IS 1140 IN REF.1, REF.3 AND REF.4, WHILE IN REF.2 IT IS 1139.
 KW Amphibian skin; Amidation.
 FT MCD_RES 12 12 AMIDATION
 SQ SEQUENCE 12 AA: 1141 MW: 2822551A3372728 CRC64;

 Query Match 1.6%; Score 4; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 82 LQCT 85
 DB 6 LQCT 9

 RESULT 40
 AUI1_LITRA
 ID AUI1_LITRA STANDARD; PRT; 13 AA.
 AC P82386;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aurein 1.1.
 OS Litoria raniformis (Southern bell frog).
 OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 OC Amphibia: Batrachia: Anura: Neobatrachia: Bufonidae; Hylidae;
 OC Pelodyadinae: Litoria.
 CX NCBI_TaxID=1116057;
 RN [1]
 RP SEQUENCE, AND FUNCTION.
 RC TISSUE=Granular dorsal gland;
 RX MEDLINE=20408845; PubMed=10951191;
 RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
 RA Wallace J.C., Tyler M.J.;
 R1 "The antibiotic and anticancer active aurein peptides from the
 RT Australian bell frogs Litoria aurea and Litoria raniformis the
 RI solution structure of aurein 1.2.";
 RL Eur. J. Biochem. 267:5330-5341(2000).
 CC [1] FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST B.CEREUS, L.LACTIS,
 CC L.INNOCUA AND S.UBERIS. PROBABLY ACTS BY DISRUPTING MEMBRANE
 CC FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
 CC [1] SUBCELLULAR LOCATION: SECRETED.

KW Antibiotic: Amidation.
 FT MOL RES 13 13
 SQ SEQUENCE 13 AA; 1447 MW; 173CB99DFRC83330 CRC64;
 Query Match 1.68; Score 4; DH 1; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.8e-03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CY 42 LFD1 45
 DB 2 LFD1 5

Search completed: April 10, 2003, 10:39:44
 Job time : 11.4706 secs

Result	No.	Score	% Match	Length	DB	ID	Description
1	474	36.9	100	2	H44549	probable ubiquitin	
2	423.5	32.2	93	2	G96570	hypothetical prote	
3	351	29.7	102	2	D84749	probable ubiquitin	
4	372	28.3	485	2	T31067	RIR repeat contain	
5	269	20.5	423	2	T15350	hypothetical prote	
6	179	13.6	172	2	A11906	ubiquitin-conjugat	
7	7	6.9	12.9	247	B42856	ubiquitin-carrier	
8	165.5	12.6	165	2	S28951	ubiquitin-conjugat	
9	164	12.5	153	2	C86304	probable ubiquitin	
10	161	12.2	153	2	S80582	hypothetical prote	
11	159	12.1	151	2	S25259	ubiquitin-conjugat	
12	159	12.1	205	2	T32959	hypothetical prote	
13	158.5	12.1	151	2	T15931	hypothetical prote	
14	157	11.9	151	2	T45200	ubiquitin-protein	
15	155	11.9	152	2	T51913	14 kDa ubiquitin c	
16	157	11.9	152	2	A24246	ubiquitin carrier	
17	157	11.9	152	2	A41222	ubiquitin-protein	
18	155.5	11.9	166	2	T43235	ubiquitin-conjugat	
19	153	11.6	152	2	A41222	ubiquitin-protein	
20	152.5	11.6	151	2	S71430	DNA repair protein	
21	151	11.5	151	2	A39392	RAD6 DNA-repair ho	
22	151	11.5	163	2	H66818	hypothetical prote	
23	151	11.5	213	1	S19158	ubiquitin-protein	
24	150	11.4	154	2	F86484	hypothetical prote	
25	148	11.3	298	2	A49630	ubiquitin conjugat	
26	145.5	11.1	199	2	T33659	hypothetical prote	
27	144.5	11.0	152	2	S43761	ubiquitin-conjugat	
28	144	11.0	148	2	T37532	ubiquitin-conjugat	
29	143.5	10.9	152	2	S43763	ubiquitin-conjugat	

```

103 98 7.5 152 2 727167
104 98 7.5 543 2 A41285
105 97 7.4 162 2 A96785
106 97 7.4 189 2 S43784
107 96.5 7.3 309 2 T34795
108 96 7.3 185 2 C84566
109 96 7.3 497 2 H42827
110 94 7.1 161 2 152053
111 94 7.1 370 2 AG0761
112 94 7.1 430 2 157013
113 92.5 7.0 170 2 A96563
114 91.5 7.0 155 2 739559
115 91 6.9 528 2 150312
116 91 6.9 2128 2 152577
117 90 6.8 496 2 T15691
118 87.5 6.7 309 2 T41494
119 87 6.5 229 2 T21439
120 87 6.6 3869 2 A49205
121 86.5 6.5 251 2 582348
122 86 6.5 647 2 A84265
123 85 6.5 139 2 T21984
124 85 6.5 214 2 T25431
125 84.5 6.4 145 2 JC5525
126 84.5 6.4 670 2 A75542
127 84 6.4 589 2 G87485
128 83.5 6.3 613 2 A35296
129 82.5 6.3 870 2 AE0208
130 82 6.2 321 2 D83204
131 81.5 6.2 1217 2 T00270
132 81 6.2 248 2 A81040
133 81 6.2 509 2 T00793
134 81 6.2 610 2 F82192
135 81 6.2 617 2 A56951

hypotheical prote
interleukin enhanc
Putative ubiquitin
ubiquitin-protein
hypotheical prote
probable ubiquitin
interleukin enhanc
ubiquitin-conjugat
guanine nucleotide
hypotheical prote
ubiquitin-conjugat
hypotheical prote
beta-spectrin - mo
hypotheical prote
zinc finger, ADP-r
hypotheical prote
All-1 protein c-ter
conserved hypothet
hypotheical prote
hypotheical prote
hypotheical prote
alpha-25-dihydro
conserved hypothet
hypotheical prote
secretogranin II p
probable outer mem
probable outer mem
hypotheical prote
conserved hypothet
hypotheical prote
ABC transporter, A
myocyte nuclear fa

```

```

DB 986 GYDKQVGTAFGRKNSIGYNTFTLNCKTMMYLMRKPPKDFELIKDHERKRGYYIKAC 1039
QY 182 ESWLETHAL 190
DB 1040 DAYMKGYLI 1048

RESULT 2
G96570
hypotheical protein F8L10.11 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96570
R:Thelodis, A.; Ecker, J.R.; Palm, C.J.; Federspiele, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.;
assen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, C.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Ye, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96570
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-543 <STO>
A:Cross-references: GB:AE005173; NID:g9454541; PIDN:AAH97864.1; GSPDB:GN00141
C:Genetics:
A:Gene: F8L10.11
A:Map position: 1

Query Match 32.24; Score 423.5; DB 2; Length 543;
Best Local Similarity 41.58; Pred. No. 3.4e-29;
Matches 85; Conservative 38; Mismatches 65; Indels 17; Gaps 4;

QY 3 LLATSLPEGIMVKTEDRMDFLSALIKGPTRTPYEDGLYLFDIQLPNLYPAVVPHECYLS 62
DB 282 ILENDLPEALSVRACESRMDLLRAVIIGAGSTPYHGLFFEDFQPTPTSPVPPVHYHS 342
QY 63 QCSG-RINPNLYDNKVCVSLGCTGKTERW-TSKSLIQVLISIOGLVNEPYNE 120
DB 342 --GGIRINPNLYNCKVCVSLGCTGKTERW-TSKSLIQVLISIOGLVNEPYNE 399
QY 121 AGFSDRGLOEGYNSRCYNEMALIRVVQSWTQVLRPPPEVEQEIRHSTGWRVLR 180
DB 400 PGYVOSAGTASGSKSVYSENVELLSEKTMVYSIRRPPOHEEYVQNHV----- 449
QY 181 IESWIFTHALLEKAQALPNGVPKAS 205
DB 450 ---FVRSHDIVKACNAYKAGAPLGS 471

RESULT 3
D84749
probable ubiquitin-conjugating enzyme E2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84749
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:1067197
A:Accession: D84749
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-182 <STO>
A:Cross-references: GB:AE002093; NID:g1707021; PIDN:AAC69130.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g33770

```


Db 2: PCVSVASPLDNNVWMMIIGADIPYEGTFRLLIFDEYENKPPHVKELSF---MF 77
 Qy 69 NPNLYNGKVCVSLGTWICKGHTERTSKSSLIQVLSIQGLITIVNEP 116
 Db 78 HPNVYANGKICUDU-----QNRWTPYDVASILTSTIOSLF--NDP 116

RESULT 7
 B42856
 ubiquitin carrier protein E2 - human
 C:Species: Homo sapiens (man)
 C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: B42856
 R:Liu, Z.; Diaz, L.A.; Haas, A.L.; Giudice, G.J.
 J. Biol. Chem. 267, 15829-15835, 1992
 A:Title: cDNA cloning of a novel human ubiquitin carrier protein. An antigenic domain sp
 this human epidermal transcript.
 A:Reference number: B42856; MUID:92348449; PMID:1379239
 A:Accession: B42856
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-247 <L1>
 A:Experimental source: keratinocyte
 A:Note: Sequence extracted from NCHI backbone (NCBIN:109695, NCHIP:109698)
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 12.9% Score 169; DB 2; Length 247;
 Best Local Similarity 24.5% Pred. No. 2.5e-07;
 Matches 60; Conservative 35; Mismatches 73; Indels 76; Gaps 10;

Qy 4 LATSFGPGIMVKTEDMDLFSALIKGPTPTPEYDGLYFDIQLDNIPYAVPPHFCYLSQ 63
 Db 45 LTADPGDGIKVFPEEDITDLOVTEEPKSTPYAGSLFRMKLLLGKGFVPSPKGYELTK 104
 Qy 64 CSGRINPNLYNGKVCVSLGTWICKGHTERTSKSSLIQVLSIQGLITIVNEP 123
 Db 105 -----HIPNVGANGKICVNL-----KRDWTAELG:RHVLIIKGLIHTPNP---ESAL 151
 Qy 124 CSGRG--LQEGYENSRYNEMALIRVVQSMTOLVRRPPEVEQELQHFSTGGWGLVNR 181
 Db 152 NEFAGKILENYE--YAAKAK-----LIE-----HGAGG----- 182
 Qy 182 ESWLHETHALPKAALNGVKKPSSPEPPAVADLSGGOOPEDGSPAPKASGSDSE 241
 Db 183 -----PSGRAENGRAIASGTASSTDPGA?-----GCP----- 210
 Qy 242 GGAQG 246
 Db 211 GGAEG 215

RESULT 8
 S28951
 ubiquitin conjugating enzyme UBC7 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YM9711.12; protein YMR022W; ub:quitin-conjugating enzyme QR18
 C:Species: Saccharomyces cerevisiae
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
 C:Accession: S28951; S29741; S54024; S29338
 R:Vassal, A.; Houlet, A.; Decoster, E.; Paye, G.
 Biochim. Biophys. Acta 1132, 211-213, 1992
 A:Title: QR18, a novel ubiquitin-conjugating enzyme in Saccharomyces cerevisiae.
 A:Reference number: S28951; MUID:93003327; PMID:1327148
 A:Accession: S28951
 A:Molecule type: DNA
 A:Residues: 1-165 <VAS>
 A:Cross-References: EMBL:X66829; NID:g4256; PIDN:CAA47302.1; PID:g4257
 R:Jungmann, J.; Reins, H.A.; Schobert, C.; Jentsch, S.
 Nature 361, 369-371, 1993
 A:Title: Resistance to cadmium mediated by ubiquitin-dependent proteolysis.
 A:Reference number: S29741; MUID:93149278; PMID:8361213
 A:Accession: S29741
 A:Molecule type: DNA

A:Residues: 1-165 <JUN>
 A:Cross-References: EMBL:X69100; NID:g5522; PIDN:CAA48846.1; PID:g5523
 R:Lye, G.; Charcher, C.M.
 submitted to the EMBL Data Library, May 1995
 A:Reference number: S54014
 A:Accession: S54024
 A:Molecule type: DNA
 A:Residues: 1-165 <LYE>
 A:Cross-References: EMBL:Z49211; NID:q798922; PID:q798933; MIPS:YMR022W
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: SGD:QR18; UBC7
 A:Cross-References: SGD:S0004624; MIPS:YMR022W
 A:Map position: 138
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 12.6% Score 165.5; DB 2; Length 165;
 Best Local Similarity 31.2% Pred. No. 2.9e-07;
 Matches 39; Conservative 28; Mismatches 45; Indels 13; Gaps 5;

Qy 9 PEGIMV-KTFEDMDLFSALIKGPTPTPEYDGLYFDIQLDNIPYAVPPHFCYLSQSGR 67
 Db 21 PGIVAGPKSENNIFWDLIQGPDPTPYADGVFNAKLEFFKDFPLSPKLIPTPSI--- 77
 Qy 68 LPNLYNGKVCVSLGT-----WICKGHTERTSKSSLIQVLSIQGLITIVNEPYNFA 121
 Db 78 LPNLYNGEVCISILHSPGDDPNMYELAFERWSPQSVKILLSVMSML--SEPNIES 134
 Qy 122 GFDSQ 126
 Db 135 GAKID 139

RESULT 9
 C86304
 probable ubiquitin-conjugating enzyme E2 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C86304
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar,
 ansen, N.F.; Hughes, R.; Huizart, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malfi, R.; Marzia
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Talio
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talio
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C86304
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-153 <SIO>
 A:Cross-References: GB:AB005172; NID:g9802775; PIDN:AAF99844.1; GSPDB:GNGG141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 12.5% Score 164; DB 2; Length 153;
 Best Local Similarity 27.6% Pred. No. 3.6e-07;
 Matches 42; Conservative 28; Mismatches 62; Indels 20; Gaps 4;

Qy 3 LLATSLPEGIMVKTEDMDLFSALIKGPTPTPEYDGLYFDIQLDNIPYAVPPHFCYLS 62
 Db 17 LISEAP-GISASPEENMYFNWMLGPTQSPYEGGVFKLEFLPEYMAAPKVRFL 75
 Qy 63 QCSGRINPNLYNGKVCVSLGTWICKGHTERTSKSSLIQVLSIQGLITIVNEPYNFA 122
 Db 76 K----YHPNCKLGRICLDL-----KDKWSPALQRTVLLSIQALSAPNP- --- 119
 Qy 123 FDSQGIQFQYENSRYNEMALIRVVQSMIQL 154

Ds 120 ---DDPLSEN:AKHWKSNFAGAVTAKWIRL 148

RESULT 10

S58092

hypothetical protein YDR092w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YD652.04

C:Species: Saccharomyces cerevisiae

C>Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999

C:Accession: S58092

R:Oliver, K.; Harris, D.

submitted to the EMBL Data Library, July 1995

A:Reference number: S58089

A:Accession: S58092

A:Molecule type: DNA

A:Residues: 1-153

A:Cross-references: EMBL:Z50112; NID:9514872; PIDN:CAA90451.1; PID:9914876; MIPS:YDR092w

A:Experimental source: strain AH972

C:Genetics:

A:Gene: SGD:URC13

A:Cross-references: SGD:S0020499; MIPS:YDR092w

A:Map position: 4R

A:Introns: 10/3

C:Superfamily: human ubiquitin-protein ligase p2

Query Match 12.2%; Score 161; DB 2; Length 153;
Best Local Similarity 29.6%; Pred. No. 6.6e-07;
Matches 37; Conservative 29; Mismatches 45; Indels 14; Gaps 4;

Qy 3 LIAISLPQIMVKTFEDRMILFSALIKGPTPEYEDGLYFDIQLPNITYPAVPPHFCYLS 62

Db 15 LVSDVFPGLAPHDNINLYFQVTEKPSQPYEDGIFUELYLIDNYPMEAPKVRFLT 73

Qy 63 QCSGRINFLYDNGKVCVSLATWICKGTERTWTSKSLQVLISIGLITL---VNEPYIN 119

Db 74 KLT---YHNIDRGRICLDVLT-----NNSPALQIRVLSIGALLASPNNDPLAN 123

Qy 120 PACPD 124

Db 124 DVAED 126

RESULT 11

S12529

ubiquitin-conjugating enzyme rhp6 - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: protein SPAC18B11.07c

C:Species: Schizosaccharomyces pombe

C>Date: 21-Nov-1993 #sequence_revision 12-Apr-1996 #text_change 10-Dec-1999

C:Accession: S12529; 137907; S54845

R:Reynolds, P.; Koken, M.H.M.; Hoeltmakers, J.H.J.; Prakash, S.; Prakash, L.

EMBL J. 9, 1423-1430, 1990

A:Title: The rhp6(+) gene of Schizosaccharomyces pombe: a structural and functional homologue of the human ubiquitin-conjugating enzyme E2

A:Reference number: S12529; MUID:90228339; PMID:2184030

A:Accession: S12529

A:Molecule type: DNA

A:Residues: 1-151 <RBY>

A:Cross-references: EMBL:Z50728; NID:9429886; PIDN:CAA90592.1; P:Z:9929893

R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, August 1995

A:Reference number: Z21753

A:Accession: 137907

A>Status: preliminary; translated from GR/EMBL/DDRC

A:Molecule type: DNA

A:Residues: 1-151 <DEV>

A:Cross-references: EMBL:Z50728; PIDN:CAA90592.1; PID:9929893; GSPDB:GN00066; SPDB:SPAC1

A:Experimental source: strain 972h; cosmid ci8811

C:Genetics:

A:Gene: rhp6

A:Map position: 1L

A:Introns: 14/3; 36/2; 76/3; 111/2

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: nucleus

Query Match 12.1%; Score 159; DB 2; Length 151;
Best Local Similarity 30.5%; Pred. No. 9.7e-07;
Matches 40; Conservative 23; Mismatches 54; Indels 14; Gaps 4;

Qy 9 PEGIMVKTFEDRMILFSALIKGPTPEYEDGLYFDIQLPNITYPAVPPHFCYLSQCSGR 68

Db 21 PAGVSGAPEDNLTWEALIFGPOETPFEDGTFLSLEFTEEPNKEPTVKFSK---MF 77

Qy 69 NPNLYDNGKVCVSLGCTWICKGTERTWTSKSLQVLISIGLITLIVNEPYNEAGFDSRG 128

Db 78 HPNVYANGELCIDIL-----QNRWSPYDVAALITISQSL--NDF--NNASPAKAE 126

Qy 129 LQEGYENSRCHY 139

Db 127 AOLHRENKEY 137

RESULT 12

T32959

hypothetical protein C35B1.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C:Accession: T32959

R:Miller, N.; Strellyes, L.; Bradshaw, H.; Keppeler, D.

submitted to the EMBL Data Library, February 1998

A:Description: The sequence of C. elegans cosmid C35B1.

A:Reference number: Z21255

A:Accession: T32959

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-205 <ML>

A:Cross-references: EMBL:AF045638; PIDN:AAC0256.1; GSPDB:GN00022; CLSP:C35B1.1

A:Experimental source: strain Bristol N2; clone C35B1

C:Genetics:

A:Gene: CESP:C35B1.1

A:Map position: 4

A:Introns: 17/3; 30/3; 160/3

C:Superfamily: yeast ubiquitin-protein ligase UBC1

Query Match 12.1%; Score 159; DB 2; Length 205;
Best Local Similarity 29.7%; Pred. No. 1.5e-06;
Matches 43; Conservative 25; Mismatches 59; Indels 18; Gaps 5;

Qy 9 PEGIMVKTFEDRMILFSALIKGPTPEYEDGLYFDIQLPNITYPAVPPHFCYLSQCSGR 68

Db 34 PAGVSGAPEDNLTWEALIFGPOETPFEDGTFLSLEFTEEPNKEPTVKFSK---MF 90

Qy 69 NPNLYDNGKVCVSLGCTWICKGTERTWTSKSLQVLISIGLITLIVNEPYNEAGFDSRG 128

Db 91 HPNVYADGSGICIDIL-----QNRWSPYDVAALITISQSL--DEPNINSPA--NSLA 139

Qy 129 LQEGYENSRCHYENALIRVYQSMQ 153

Db 140 AOLYQENRRREYK-----RVQGIHQ 160

RESULT 13

T51931

hypothetical protein NhrAD6 [imported]; - Haematococcus haematococcus

C:Species: Haematococcus haematococcus

C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000

C:Accession: T51931

R:Wu, Q.; Aist, J.R.; Wirsal, S.G.; Turgeon, B.G.; Yoder, O.C.; Sandrock, T.

submitted to the EMBL Data Library, January 1997

A:Description: Nectria haematococca mating population VI NhrAD6 and NhrK1 genes.

A:Reference number: Z25871

A:Accession: T51931

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-151 <WUQ>

A:Cross-references: EMBL:U86521; PIDN:AAB47850.1

A:Experimental source: strain T213 mating population VI

C:Genetics:

A:Cross-references: EMBL:U04306; NID:g476115; PIDN:AAH60669.1; PID:g476117
C:Genetics
A:Introns: 15/2; 42/2; 51/1; 81/1; 110/3
C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.9%; Score 157; DB 2: Length 152;
Best Local Similarity 30.3%; Pred. No. 1.5e-06;
Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;

Oy 9 PEGIMVKTFEDRMOLFSAALIKGTRTPYEDGLYLFDIQLNIIYVAPPPHFCYLSQCSGRL 68
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 21 PQCVGASPSNNIMQWNAVIFGEGTFEDGTGFKLVIEFSEFPYKPPVTRFLSK--MF 77
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 69 NPNLYDNKVCVSLGWTGWICKGTRWTSKSSLLQVLISIOGLILVNEPYNFAGFDSRG 128
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 78 HPNVAADGSICLDIL-----QNRWSPDYDVSLLTSIQSL--DEPNPNSPA--NSQA 126
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 129 LQGYENSGCYNEMALIRVQVS 150
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 127 AOLYQNRKREYFKRVSIVPEQS 148
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
RESULT 16
A42416
ubiquitin carrier protein E2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A42416
R:Wang, S.S.; Dumas, F.; Banville, D.
J. Biol. Chem. 267, 6495-6501, 1992
A:Title: A rabbit reticulocyte ubiquitin carrier protein that supports ubiquitin-dep
A:Reference number: A42416; MUID:92202189; PMID:1313008
A:Accession: A42416
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-52 <WIN>
A:Cross-references: GB:M62387; NID:g165779; PIDN:AAA1492.1; PID:g165780
A:Experimental source: reticulocyte
A:Note: sequence extracted from NCBI backbone (NCBIN:89551, NCBI:P:89553)
C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.9%; Score 157; DB 2: Length 152;
Best Local Similarity 30.3%; Pred. No. 1.5e-06;
Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;

Oy 9 PEGIMVKTFEDRMOLFSAALIKGTRTPYEDGLYLFDIQLNIIYVAPPPHFCYLSQCSGRL 66
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 21 PQCVGASPSNNIMQWNAVIFGEGTFEDGTGFKLVIEFSEFPYKPPVTRFLSK--MF 77
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 69 NPNLYDNKVCVSLGWTGWICKGTRWTSKSSLLQVLISIOGLILVNEPYNFAGFDSRG 128
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 78 HPNVAADGSICLDIL-----QNRWSPDYDVSLLTSIQSL--DEPNPNSPA--NSQA 126
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Oy 129 LQGYENSGCYNEMALIRVQVS 150
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 127 AOLYQNRKREYFKRVSIVPEQS 148
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
RESULT 17
B41222
ubiquitin-protein ligase (EC 6.3.2.19) E2b - human
N:Alternate names: ubiquitin carrier protein E2; ubiquitin-conjugating enzyme HHA6B
C:Species: Homo sapiens (man)
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 03-Jun-2002
C:Accession: B41222; S12530; S17516
R:Koken, M.H.M.; Reynolds, P.; Jaspers-Dekker, I.; Prakash, S.; Bootsma
Proc. Natl. Acad. Sci. U.S.A. 88, 8865-8869, 1991
A:Title: Structural and functional conservation of two human homologs of the yeast E2
A:Reference number: A41222; MUID:92020951; PMID:1717990
A:Accession: B41222
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-152 <KOK>

A:Cross-references: GB:M74525; NID:g184045; PIDN:AAA35982.1; PID:g184046
R.Schneider, K.; Ekerskorn, C.; Lottspeich, F.; Schweiger, M.
EMBO J. 9, 1431-1435, 1990

A>Title: The human ubiquitin carrier protein E2(Mir) - 17000 is homologous to the yeast
E2(Mir)

A:Reference number: S12530; WUID:90228340; PMID:2158443

A:Accession: S12530

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-152 <SCH>

A:Cross-references: GB:X53251; NID:g30953; PIDN:CAA37339.1; PID:g30954
R.Wollfendin, C.; Chen, Z.; Staskus, K.; Retzel, E.F.; Plagemann, P.G.W.
Biochim. Biophys. Acta 1090, 81-85, 1992

A>Title: Mammalian mRNAs encoding protein closely related to ubiquitin-conjugating enzyme
E2

A:Reference number: S17516; MUIP:91355233; PMID:1883845

A:Accession: S17516

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-21, C, 24-40, I, 42-53, R, 55-152 <WOF>

C:Genetics:

A:Gene: CDB:UBH2H; UNC2; HHR6A

A:Cross-references: GDB:128627; OMIM:179095

A:Map position: 6q23-q31

C:Superfamily: human ubiquitin protein ligase E2

C:Keywords: ligase; protein degradation; ubiquitination

F:88/Active site: Cys 4status predicted

Query Match 11.9%; Score 157; DB 2; Length 152;
Best Local Similarity 30.3%; Pred. No. 1.5e+06;
Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;

QY 9 PGICMWKIFEDRMDFLSALIKGPTPTTPYDGLYLFDIQLPNIIYPVAPHPFCYLSQCGR 68
| | :
DE 21 PCVGSGAPSENMMORNAVIFGFETPDGTETKLVIREFSEEPKNKPPTVFRLSK---MF 77
| | :
QY 69 PNLIYDKVCVSLGLTGWTGWIGKTERTWSKSSLLQVLISIOGLLVNEPYNEAGFDSDRG 128
| | :
Db 78 HNIVYACVSICLDL-----QNWSPTYDVSVLTISQSIL--DEPNMNSPA--NSQA 126
| | :
QY 129 LOAYENSRCYNEMAIRVYQS 150
| | :
Db 127 ASLYCENKREYKRVSAIVEQS 148
| | :

RFSLL 16
T43235
ubiquitin-conjugating enzyme ubcP3 - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: UbcP3
C:Species: Schizosaccharomyces pombe
C:date: 11-Jan-2000 *sequence_revision 11-Jan-2000 *text_change 21-Jul-2000
C:Accession: T43235
K:Osaka, F.; Seino, H.; Seno, T.; Yamao, F.
Mol. Cell. Biol. 17, 3388-3397, 1997

A>Title: An ubiquitin-conjugating enzyme in fission yeast that is essential for the onset
of meiosis

A:Reference number: Z22356; MUIP:97295689; PMID:9154838

A:Accession: T43235

A>Status: preliminary; translocated from GB/FMH1/DDHJ

A:Molecule type: mRNA

A:Residues: 1-166 <OSA>

A:Cross-references: EMBL:D85544; NID:g2190251; PIDN:BAA20373.1; PTD:g2190252

C:Genetics:

A:Gene: ubcP3

C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.9%; Score 156.5; DB 2; Length 166;
Best Local Similarity 32.0%; Pred. No. 1.8e+06;
Matches 40; Conservative 27; Mismatches 45; Indels 13; Gaps 6;

QY 9 PGICMW-KTFEDRMDFLSALIKGPTPTTPYDGLYLFDIQLPNIIYPVAPHPFCYLSQCGR 67
| | :
Db 22 PDGITAGESNEDOFFTWDLCLIQSDPTPEGLYPAILAFPSDPLGPILKPF--ECE-F 78
| | :
QY 68 LNPNLYNGKCVSLEI-----GIWKDGTERNTSKSLIQVLSIQGLLVNEPYNEFA 121
| | :

QY 4 LATSLEPGIMVKTEDMDLFSALIKGPIRTPEYEGLYLFDIQLPNIPYPAVPPHFCYLSQ 63
 Db 16 MQICRPAGVSPDNDVMTNAVITGPADIPPEGDTFLVMHFEEQYIPKPKPSVAFISE 75
 QY 64 CSRLNPN:YUNGKVCVSLICW:GKGTERTWSKSS:LOVLIS:OGILLVNEPYNEAGE 123
 Db 76 ---MFIPNYYATGELCGLDIL-----QNRWSPYTDVAANV:SIQSLL--NDP----- 116
 QY 124 DSDRGLOQCYENSKYCNK 141
 Db 117 --NTGSPANVEASN:YKD 132

RESULT 21
 A39392
 RAD6 DNA-repair homolog chr6 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 20-Sep-1999
 C:Accession: A39392
 R:Koken, M.; Reynolds, P.; Bootsma, D.; Hoelmakers, J.; Prakash, S.; Prakash, L.
 Proc. Natl. Acad. Sci. U.S.A. 88, 3832-3836, 1991
 A:Title: chr6, a Drosophila homolog of the yeast DNA-repair gene RAD6.
 A:Reference number: A39392; MUID:91219466; PMID:1902572
 A:Accession: A39392
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-151 <KOK>
 A:Cross-references: GB:M63791; GH:M63792; NID:g156731; PID:g156732
 C:Genetics:
 A:Gene: Flybase:fbcd6
 A:Cross-references: Flybase:Fgn0004436
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.5%; Score 151; DB 2; Length 151;
 Best Local Similarity 25.2%; Pred. No. 4.9e-06;
 Matches 40; Conservative 30; Mismatches 58; Indels 14; Gaps 4;

QY 9 PEIMVKIFEDMDLFSALIKGPIRTPEYEGLYLFDIQLPNIPYPAVPPHFCY:SQCSGRL 68
 Db 21 PIGVSCAPTDNNIMWNAVIFGPHDIPFEDGTFTKITEETEPNKPOTVREYSKY---F 77
 QY 69 NPN:YUNGKVCVSLIGTWIGKGTERTWSKSSLLQVLISIQGLLVNEPYNEAGFSDRG 128
 Db 78 HPNVYADGKGLDIL-----QNRWSPYDVSALTSIQSL--SDPNNSPANSTAAQ 128
 QY 129 LQGYENSKYCNEMALIRVQS 150
 Db 129 LYK--ENRVEYKRVKAVRQS 148

RESULT 22
 B96818
 hypothetical protein F9K20.8 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: B96818
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chn, C.W.; Chang, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ausen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: B96818; MUID:21016719; PMID:11130712
 A:Accession: B96818
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-163 <STO>
 A:Cross-references: GB:AE005173; NID:g3834310; PID:AA083026.1; GSPDB:GN00141
 C:Genetics:

A:Gene: F9K20.8
 A:Map position: 1
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.5%; Score 151; DB 2; Length 163;
 Best Local Similarity 31.3%; Pred. No. 5.4e-06;
 Matches 36; Conservative 24; Mismatches 43; Indels 12; Gaps 4;

QY 3 LATSLEPGIMVKTEDMDLFSALIKGPIRTPEYEGLYLFDIQLPNIPYPAVPPHFCYLS 62
 Db 17 LLSHPAP-GISASPSKNNRYFNWILGPTQSPYEGVFKIELFLPEHYPMAPKVRFLIT 75
 QY 63 QCSGRNPNLYDNGKVCVSLIGTWIGKGTERTWSKSSLLQVLISIQGLLVNEP 116
 Db 76 K-----YHPNIDKIGRICDIL-----KDKWSPALQTRTVLLKSIQALISAPNP 120

RESULT 23
 S19158
 ubiquitin-protein ligase (EC 6.3.2.19) E2 - African swine fever virus (isolate Malaw
 C:Species: African swine fever virus, ASFV
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 03-Jun-2002
 C:Accession: S19158
 R:Ringamp, P.M.; Arnold, J.E.; Mayer, R.J.; Dixon, L.K.
 EMBO J. 11, 361-366, 1992
 A:Title: A ubiquitin conjugating enzyme encoded by African swine fever virus.
 A:Reference number: S19158; MUID:92155177; PMID:1310934
 A:Accession: S19158
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-213 <HIN>
 A:Cross-references: EMBL:X62440; NID:g58649; PID:CAA44305.1; PID:g58650
 C:Superfamily: African swine fever virus ubiquitin-protein ligase E2
 C:Keywords: ligase; protein degradation

Query Match 11.5%; Score 151; DB 1; Length 213;
 Best Local Similarity 32.4%; Pred. No. 7.9e-06;
 Matches 36; Conservative 20; Mismatches 43; Indels 12; Gaps 4;

QY 10 EGTWVATFEDMDLFSALIKGPIRTPEYEGLYLFDIQLPNIPYPAVPPHFCYLSQCSGR 69
 Db 19 EHFKISVNEEDNTEMDVLKGPDTLYEGGLFKAKIVFPKPYEPPEPLITFSE---MWH 75
 QY 70 PNYDNGKVCVSLIGTWIGKGTERTWSKSSLLQVLISIQGLLVNEP 116
 Db 76 PNIYSDCKICISTILH---GDNAEEQGMWSPACKIDTVLISV--ISLINHP 121

RESULT 24
 E86484
 hypothetical protein F7F23.6 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: E86484
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon,
 ausen, N.F.; Chang, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: E86484; MUID:21016719; PMID:11130712
 A:Accession: E86484
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-154 <STO>
 A:Cross-references: GB:AF005172; NID:g6978907; PID:AAF34303.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.4%; Score 150; DB 2; Length 154;
 Best Local Similarity 33.3%; Pred. No. 5, 1e-06;
 Matches 32; Conservative 23; Mismatches 27; Indels 12; Gaps 3;

QY 24 FSALIKGPTPIYKNGLYLDFDIOLPNLYPAVPPHFCYLSQCSGRHNPCLYDNCKVCVSL 83
 DB 39 WIAVINGDGTPIYEGGMFN:SIKFP:YDFPKPKFTKPTI---YHPNINDRGSGCMNII 95

QY 84 GTWICKGTERWTSKSSLLQVLIISLOGLIYNRP 116
 DB 96 -----KUKWTPA:MVEKVLISF---LILKPKP 119

RESULT 25
 A49630
 ubiquitin conjugating enzyme - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 05-Nov-1999
 R:Pilon, S.E.; Leppik, K.A.; Do, H.N.; Groudine, M.
 Proc. Natl. Acad. Sci. U.S.A. 90, 10484-10488, 1993
 A:Title: Cloning of the human homolog of the CDC34 cell cycle gene by complementation in
 A:Reference number: A49630; MUID:94068425; PMID:8248134
 A:Accession: A49630
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1 298 <RES>
 A:Cross-references: GB:L22005; NID:q388308; PIDN:MACJ7534.1; FID:q388309
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 31.3%; Score 148; DB 2; Length 298;
 Best Local Similarity 30.3%; Pred. No. 2, 3e-05;
 Matches 44; Conservative 28; Mismatches 55; Indels 18; Gaps 7;

QY 10 ECLMKTFEDRMFLFS--ALLKGP:HTPYEDGLIFDIOLPNLYPAVPPHFCYLSQCSGR 67
 DB 88 EGRV-FLVDRGRLYNNEVALFQPPNYYEGYFKAK:KFPIDYSPAPAFRLIK--M 143

QY 68 LHNLYDNCKVCVSLGTWTKGK-----GIERWTSKSSLLQVLIISLOGLIYNRP 120
 DB 144 WHPNLYE:SDVCISLILPDPVQSGSLPSRWNPUNVTRTILSV--ISLNEPNTFSP 202

QY 121 AGFESD---RC:GRGYNSRCYNEM 142
 DB 202 ANVDASVWYKWKESK:KDRFY:DI 225

RESULT 26
 T33629
 hypothetical protein F40G9.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T33629
 R:Graves, T.; Sutterer, C.; Ozersky, P.
 submitted to the EMBL Data Library, October 1998
 A:Description: The sequence of C. elegans cosmid F40G9.
 A:Reference number: Z21378
 A:Accession: T33629
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-199 <GRA>
 A:Cross-references: EMBL:AF099919; PIDN:MAC68796.1; GSPDB:GN00021; CESP:F40G9.3
 A:Experimental source: strain Bristol N2; clone F40G9
 C:Genetics:
 A:Gene: CESP:F40G9.3
 A:Map position: 3
 A:Introns: 40/3; 72/3; 100/2; 142/3
 C:Superfamily: yeast ubiquitin-protein ligase UBC1

Query Match 11.1%; Score 145.5; DB 2; Length 199;
 Best Local Similarity 29.9%; Pred. No. 2, 2e-05;
 Matches 32; Conservative 25; Mismatches 39; Indels 11; Gaps 3;

QY 11 GIMVKTEDRMFLFSALIKGPTPIYKNGLYLDFDIOLPNLYPAVPPHFCYLSQCSGR 70
 DB 26 GIMIELNENLEIKGHTRGPDTPYAGGMFDLDIKIPDQYFPSPNVKFSIKI---WHP 82

QY 71 NL-YDNCKVCVSLGTWTKGTERWTSKSSLLQVLIISLOGLIYNRP 116
 DB 83 NVSQTGVICLDIL-----KQWAASTLRTVLVLSIALCMCTPEP 122

RESULT 27
 S43781
 ubiquitin-conjugating enzyme UBC1 - Arabidopsis thaliana
 N:Alternate names: ubiquitin-conjugating enzyme E2
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
 C:Accession: S43781
 R:Sullivan, M.F.; Carpenter, T.B.; Vierstra, R.D.
 Plant Mol. Biol. 24, 651-661, 1994
 A:Title: Homologues of wheat ubiquitin-conjugating enzymes - taUBC1 and taUBC4 are c
 A:Reference number: S43781; MUID:94207190; PMID:8155884
 A:Accession: S43781
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-152 <SUL>
 A:Cross-references: EMBL:L19351; NID:g431259; PIDN:AAA32897.1; PID:g431260
 C:Genetics:
 A:Gene: UBC1
 A:Introns: 42/2; 51/1; 81/1; 110/3
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.0%; Score 144.5; DB 2; Length 152;
 Best Local Similarity 26.4%; Pred. No. 1, 8e-05;
 Matches 42; Conservative 29; Mismatches 53; Indels 35; Gaps 5;

QY 9 PEGIMVKTFEDRMFLFSALIKGPTPIYKNGLYLDFDIOLPNLYPAVPPHFCYLSQCSGR 68
 DB 21 PACISGAPODNNIMLWNAVIFGPDTPWGGTKLSLOFSELYPNKPTVRFVSR---MF 77

QY 69 NPCLYDNCKVCVSLGTWTKGTERWTSKSSLLQVLIISLOGLIYNRP 125
 DB 78 HPNTYADGSGICLDIL-----QNMSPYDVAALISIQSLICDPNPNSPANSFA---- 126

QY 126 DRGLQCYNSRCYNEMALRVQSMFTQLVRRPPEVEQ 164
 DB 127 -----ARWYSE-----SKREYNRRVRDVEQ 147

RESULT 28
 T37532
 ubiquitin-conjugating enzyme - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T37532
 R:Murphy, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z21721
 A:Accession: T37532
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-148 <MUR>
 A:Cross-references: EMBL:Z98595; PIDN:CAH11181.1; GSPDB:GN00066; SPDB:SPAC111E3.04C
 A:Experimental source: strain 972h; cosmid cl1E3
 C:Genetics:
 A:Gene: SPDB:SPAC111E3.04C
 A:Map position: 1
 A:Introns: 9/3; 63/3; 92/1
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.0%; Score 144; DB 2; Length 148;
 Best Local Similarity 29.6%; Pred. No. 2e-05;
 Matches 32; Conservative 25; Mismatches 41; Indels 10; Gaps 2;

```

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #Text_change 20-Jun-2000
C/Accession: T18512
R/Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A/Reference number: Z18935
A/Accession: T18512
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-157 <IAB>
A/Cross-references: EMBL:Z98551; PDB:CAH11153.1
C/Genetics:
A/Map position: 3
A/Note: C0855W
C/Superfamily: human ubiquitin-protein ligase E2

Query Match      10.7%; Score 141; DB 2; Length 157;
Best Local Similarity 27.7%; Pred. No. 3.9e-05;
Matches 36; Conservative 25; Mismatches 55; Indels 14; Gaps 4;

QY 17 FFDMDLSALIKGPTRYEDGLYFDIQLPNIPYAPVPPHYFCYLSQSGRINPNLYDNG 76
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 29 FADNTMYCHAIIRPDITWEGGIFHLITHESEYPVSPPKIRFLSKI--YHNIVSDG 85

QY 77 KYCVSLGTWIGKGTERTWSKSLQVLISITGLLWNEPYYNEAGFSDGLOEGTENS 136
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 86 NCLDLI-----QNWSPYDITSLISIQSL--NDP--NTSSPANPEAKRIFINNR 134

QY 137 RCYNEMALIR 146
   |||: |||
Db 135 NLYNRVLVR 144

RESULT 31
UOXPAS
ubiquitin-protein ligase (EC 6.3.2.19) E2 - African swine fever virus (Strain RA71V)
C/Species: African Swine fever virus, ASFV
C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #Text_change 03-Jun-2002
C/Accession: F39448
R/Rodriguez, J.M.; Salas, M.L.; Vinuela, E.
Virology 186, 40-52, 1992
A/Title: Genes homologous to ubiquitin-conjugating proteins and eukaryotic transcrip
A/Reference number: A39448; MUID:92087485; PMID:1309282
A/Accession: F39448
A/Molecule type: DNA
A/Residues: 1-215 <R0D>
A/Cross-references: GB:M77121; NID:q210618; PDB:AAA42704.1; PTD:q210624
C/Superfamily: African swine fever virus ubiquitin-protein ligase E2
C/Keywords: ligase; protein degradation

Query Match      10.6%; Score 140; DB 1; Length 215;
Best Local Similarity 27.9%; Pred. No. 7.4e-05;
Matches 31; Conservative 25; Mismatches 43; Indels 12; Gaps 4;

QY 10 EGIMVKTEDRMDFLSNLKGTTRTPYEDGLYLFEDIQLPNIYPVPPHYFCYLSQSGRLN 69
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 19 ENKISVNNITMDWILRGPDPTLYEGGFLKAKVAFPPPYAPPKLITFSE---MWH 75

QY 70 PNIYDNGKVCVSLIGTWCKGTR---WTSKSLIQVLISITGLLWNEP 116
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 76 PNIYDGRGLCSILH---GDNAEEGQMTWSPAQKIDTILSV--ISLINEP 121

RESULT 32
JC4894
ubiquitin-protein ligase (EC 6.3.2.19) E2N - human
N/Alternate names: ubiquitin carrier protein E2; ubiquitin-conjugating enzyme E2N (U
C/Species: Homo sapiens (man)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #Text_change 03-Jun-2002
C/Accession: JC4894
R/Yamaguchi, I.; Kim, N.S.; Sokine, S.; Seino, H.; Osaka, F.; Yamao, F.; Kato, S.
J. Biochem. 120, 494-497, 1996
A/Title: Cloning and expression of cDNA encoding a human ubiquitin-conjugating enzym
A/Reference number: JC4894; MUID:97058291; PMID:8902611

```

A:Accession: J04894

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-152 <YAM>

A:Cross-references: DBJ:083004; NID:q1181557; P1DN:BAA1675.1; PID:q1181558

C:Genetics:

A:Gene: GDB:UB82N

A:Cross-references: GDB:6053724

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: ligase; protein degradation; ubiquitination

F.87/Active site: cys *status predicted

Query Match 10.4%; Score 137; DB 2; Length 152;

Best Local Similarity 24.5%; Pred. No. 8.4e-05;

Matches 35; Conservative 3; Mismatches 63; Indels 14; Gaps 4;

QY 3 ILATSLPEGIKVTEDRMOLFSAIKGPTKPYEDGILYFDIQLPNTYPAVPPHFCYLS 62

DB 15 LLAHPYV-GIKAHPDFSNARFHVVIAGPDSPFGCTPK-KIFLPEYPMAAKPKVRFMT 73

QY 63 QSGRLNPILYDNGKVCVSLGATWIGKTERWTSKSLQVLISQGLIIV---NEPYYN 119

DB 74 KLVVHPVNDLGRICLDIL-----KDKWSPALQIRIVLLSTQALISAPNDDPLAN 123

QY 120 KACFDSRGLQCYNSRCYNEM 142

DB 124 DVAEQWKTNEAQAETETARATRL 146

RESULT 35

G90107

ubiquitin-conjugating enzyme E2-21 KD [imported] - Guillardia theta nucleomorph

C:Species: nucleomorph Guillardia theta

A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C:Date: 10-May-2001 *sequence_revision 10-May-2001 *text_change 15-Jun-2001

C:Accession: G90107

R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rea

Nature 410, 1091-1096, 2001

A:Title: The highly reduced genome of an enslaved algal nucleus.

A:Reference number: A59082; MUID:11323671; PMID:11323671

A:Accession: G90107

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-144 <DGD>

A:Cross-references: GB:AJ010592; NID:q12506059; P1DN:CAC27017.1; GSPDB:GN00151

C:Genetics:

A:Map position: 2

A:Gene: nucleomorph

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: nucleomorph

Query Match 10.3%; Score 136; DB 2; Length 144;

Best Local Similarity 29.4%; Pred. No. 9.5e-05;

Matches 35; Conservative 21; Mismatches 41; Indels 22; Gaps 6;

QY 12 IMVKTEDRMDFSAIKGPTKPYEDGILYFDIQLPNTYPAVPPHFCYLSQSGRLNPN 71

DB 22 INIKPYEDLLKWKGFIIIGNPTFYCKSPNIECSVPLSPKTFVDTI---FHPN 78

QY 72 LY-DNKKVCVSLG-ATWICKTERWTSKSLQVLISQGLIIV-----NEPYYNAG 122

DB 74 VYPSNPGICLDILIKNQW---[PAWT]-----ILFSCQALIVLLTNPEPNSP-NCNAG 126

RESULT 34

S32672

ubiquitin protein ligase (EC 6.3.2.19) UBC10 - Arabidopsis thaliana

N:Alternate names: ubiquitin-conjugating enzyme UBC10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 13-Jan-1995 *sequence_revision 13-Jan-1995 *text_change 03-Jun-2002

C:Accession: S32672

R:Girod, P.; Carpenter, T.H.; van Nocker, S.; Sullivan, M.; Vierstra, R.D.

submitted to the EMBL Data Library, August 1992

A:Description: Homologs of the essential ubiquitin conjugating enzymes ubc1, 4, and

A:Reference number: S32672

A:Accession: S32672

A:Molecule type: mRNA

A:Residues: 1-148 <GIR>

A:Cross-references: FMB:0214991; NID:q297877; P1DN:CAA76715.1; PID:q297878

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: ligase

Query Match 10.3%; Score 135; DB 2; Length 148;

Best Local Similarity 30.9%; Pred. No. 0.00012;

Matches 30; Conservative 20; Mismatches 35; Indels 12; Gaps 3;

QY 22 DLF--SALIKGP-KTPYEDGILYFDIQLPNTYPAVPPHFCYLSQSGRLNPNLYDNCKVC 79

DB 29 DMFHQWATINGPSESYPAGVFLVTHFFPDYPPKPKVAFRIKV---FHPNINSNSIC 85

QY 80 VSLIGTGWIGKTERWTSKSLQVLISQGLIIVNEP 116

DB 86 LDHL-----KQWSPALITISKVLSTCSLGDPNP 115

RESULT 35

T08465

ubiquitin-protein ligase (EC 6.3.2.19) - fruit fly (Drosophila melanogaster)

N:Alternate names: ubiquitin conjugating enzyme

C:Species: Drosophila melanogaster

C:Date: 11-Jun-1999 *sequence_revision 11-Jun-1999 *text_change 03-Jun-2002

C:Accession: T08465

R:Kirby, R.J.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z16421

A:Accession: T08465

A:Status: preliminary; translated from GB/EMBL/UDBAJ

A:Molecule type: mRNA

A:Residues: 1-199 <KIR>

A:Cross-references: EMBL:X92838; NID:q1359613; PID:g1354722

A:Experimental source: strain Canton-S

C:Genetics:

A:Gene: UbcD4

A:Cross-references: FlyBase:FBgn0015321

C:Superfamily: yeast ubiquitin-protein ligase UBC1

C:Keywords: cell cycle control; DNA repair; ligase; protein degradation

Query Match 10.3%; Score 135; DB 2; Length 199;

Best Local Similarity 23.4%; Pred. No. 0.00018;

Matches 44; Conservative 26; Mismatches 82; Indels 36; Gaps 6;

QY 12 IMVKTEDRMDFSAIKGPTKTPYEDGILYFDIQLPNTYPAVPPHFCYLSQSGRLNPN 71

DB 27 IKIELVNDSTELRGEIAGPDPTPYEGGKFLVKVPEIYPPENPKAWFITRI---WHPN 83

QY 72 LYD-NGKVCVSLIGTWIGKTERWTSKSLQVLISQGLIIVNEPYNEAGFDSRGLQ 130

DB 84 ISSVGTGAIICLDIL-----KDNWAAAMTRIV:LISQALAAAEF-----DDPDQAV 129

QY 131 EGYENSRCSYEMALIRWOSMTOLVRRPPEVE-----EQEIRQHESTGG 174

DB 130 VAYGPKDYDLDFLL---TAKIWTNAYAGPHITFPDCKSKTQRLRMGIDEHEARAVLSKEN 187

QY 175 WRIVNRIE 182

DB 188 WMLEKATE 195

RESULT 36

A40797

ubiquitin-conjugating enzyme - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 17-Jul-1992 *sequence_revision 17-Jul-1992 *text_change 16-Jul-1999

C:Accession: A40797

R:Chen, Z.; Niles, E.G.; Pickart, C.M.

J. Biol. Chem. 266, 15658-15704, 1991

Plant Mol. Biol. 24, 651-661, 1994
 A:Title: Homologues of wheat ubiquitin-conjugating enzymes - TaUBC1 and TaUBC4 are encoded by two genes
 A:Reference number: S43781; MUID:94207190; PMID:8155884
 A:Accession: S43782
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-150 <SCL>
 A:Cross-references: EMBL:L19352; NID:q431261; PID:AAA32898.1; PID:q431262
 C:Genetics:
 A:Gene: UBC3
 A:Introns: 42/2; 51/1; 81/1; 110/3
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 10.2% Score 134; DB 2; Length 150;
 Best Local Similarity 27.6% Prod. No. 0.00015;
 Matches 40; Conservative 24; Mismatches 61; Indels 20; Gaps 4;

Qy 9 PEGIMKTYEDHMLFSLKNGTIRPYEDGLYFDIGIPNIYPAVDPHECYLSQCSGRL 68
 Db 21 PVGISGAPQDNNIMHNNALTFPED:PWDGG:FKLTILFTEDYENKPPIVREYSR---MF 77
 Qy 69 NPNLYDN:KVCVSLIGTWISKGTFRWTSKSLQVLIISIQGLILVNRPPYNFAGEDSDRG 128
 Db 78 HPNIYADGSICLLEI-----QWNSPIYDVANVLISQSLCDPNP-----DSPAN 123
 Qy 129 LQES---YENSRCYNEMALIRVWS 150
 Db 124 AEAARLFSNKREYNKRVIRIVEQS 148

Search completed: April 10, 2003, 10:33:41
 Job time : 17.6863 secs

Genome version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:27:50 : Search time 43.2353 Seconds
(without alignments)
770.497 Million cell updates/sec

Title: US-09-930-026-1

Perfect score: 1415

Sequence: 1 MALLATSLPEGIMVK:FEGR.....PQFASGGSYSAGAG:AFS 250

Scoring table: BLASTM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB Seq length: 2000000000

Post-processing: Minimum Match 0%

Listing First 135 summaries

Database : A:Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1315	100.0	250	23	AAU97494 Human ubiquitin-co
2	1291	98.9	466	23	AAU94354 Human protein: sequ
3	1242	94.4	300	23	AAU90275 Human polypeptide
4	551.5	41.9	1398	22	AAU64569 Drosophila melanog
5	372.5	28.3	4904	22	AAU62249 Drosophila melanog
6	370	28.2	527	22	AAU60181 Human protein: sequ
7	370	28.2	1648	22	AAU63174 Human ORFX ORF2938
8	370	28.1	1667	22	AAU95564 Human protein: sequ
9	370	28.1	4623	22	AAU97833 Human: apoptosis-in
10	338	25.7	156	21	AAU27993 Human: secreted pro

Human protein sequ
Human madt protein
Human protein sequ
Human secreted pro
Human ubiquitin-co
Protein E2-Epf dif
Human cancer assoc
Human colon cancer
Human ovarian anti
Arabidopsis thalia
Arabidopsis thalia
S cerevisiae apopt
Arabidopsis thalia
Human ubiquitin-li
Human polypeptide
Human ubiquitin-co
Lung tumour protei
Novel human enzyme
Human polypeptide
Human ovarian anti
Zea mays protein f
Zea mays protein f
Zea mays Rad6 prot
Drosophila melanog
Human E2 protein
Human rad6 homolog
Human rad6 homolog
Human ovarian anti
Human cancer assoc
Human colon cancer
Breast and ovarian
Human secreted pro
Human secreted pro
Novel human protein
Zea mays protein f
Zea mays Rad6 prot
Arabidopsis thalia
Zea mays protein f
Zea mays protein f
Arabidopsis thalia
Drosophila melanog
Human UBC/CDC34 pr
Human UBC/CDC34 p
Human UBC/CDC34
Protein CDC34 diff
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
C albicans apoptos
Human ubiquitin-co
Zea mays Rad6 prot
Human ovarian anti
Drosophila melanog
Arabidopsis thalia
Ubiquitin-conjugat
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Cotton plant ubiq
Arabidopsis thalia
Zea mays protein f
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Zea mays protein f
Human ubiquitin co
Human ubiquitin-co
Amino acid sequenc

XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0709253.
PR 11 JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RUS :NSL.
XX
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WF: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PI of an oligonucleotide comprising a sequence complementary to the
PT full-length cDNAs.
XX
XX Claim 8; SEQ ID 14875; 2537pp - CD KOM: English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH0166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB32446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 466 AA;
Query Match: 98.5%; Score 1301; DB 22; Length 466;
Best Local Similarity 99.6%; Pred. No. 1 6e-124;
Matches 247; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALLATSLPEGIMVKYTFEDRMDFLSALIKGPTPTPYEDGGLYLFDIQLPNYPAVPPHFCY 60
DB 136 MALLATSLPEGIMVKYTFEDRMDFLSALIKGPTPTPYEDGGLYLFDIQLPNYPAVPPHFCY 195
QY 61 LSQSGRINPNLNDKVKCVSLGGTWICKGIERWTSSKSLQLQVLISIQGLILVNEPPYNE 120
DB 196 LSQSGRINPNLNDKVKCVSLGGTWICKGIERWTSSKSLQLQVLISIQGLILVNEPPYNE 255
QY 121 AGFSDRLGLQEGYNSRCYNEMALIRVQSMTOIVRRPPEVFEQTRQIFSTGGWRVLNVR 180
DB 256 AGFSDRLGLQEGYNSRCYNEMALIRVQSMTOIVRRPPEVFEQTRQIFSTGGWRVLNVR 315
QY 181 IESWLETHALLEKAQALPNKVEKSPPEPPAVAELSNSGQQQEPEDGPGAPCEASQSDS 240
DB 316 IESWLETHALLEKAQALPNKVEKSPPEPPAVAELSNSGQQQEPEDGPGAPCEASQSDS 375
QY 241 ECGAAGLA 248

DB 376 ECGAAGLA 383
||||-||
RESULT 3
ABR90275
ID ABR90275 standard; Protein; 300 AA.
XX
AC ABR90275;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2651.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN W020019C304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
DR WPI: 2002-122018/16.
DR N-PSDB; ABL90684.
XX
PI Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PI prevention of neural, immune system, muscular, reproductive,
PI gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PI disorders.
XX
PS Claim 11; SEQ ID NO 2651; 2081pp + Sequence listing; English.
XX
XX The invention relates to novel genes (ABL90449-ABL90853) and proteins
CC (ABR9040-ABR90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC the nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 300 AA;
Query Match 94.4%; Score 1242; DB 23; Length 300;
Best Local Similarity 99.6%; Pred. No. 9.6e-119;
Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 13 MYKTFEDRMDFLSALIKGPTPTPYEDGGLYLFDIQLPNYPAVPPHFCYLSQSGRLNPNL 72
DB 1 MYKTFEDRMDFLSALIKGPTPTPYEDGGLYLFDIQLPNYPAVPPHFCYLSQSGRLNPNL 60

QY 73 YDNKVCVSLGCTWAKGTEHWTSKSLLOVLSISQGLLVNRPYYNEAGFSDRGLEG 132
 DB 61 YDNKVCVSLGCTWAKGTEHWTSKSLLOVLSISQGLLVNRPYYNEAGFSDRGLEG 120
 QY 133 YNSRCYNEMALIRVQSM:QIVRRPPEVFEQ:HQHSTGGWRLVNRLESWLETHALLE 192
 DB 121 YNSRCYNEMALIRVQSM:QIVRRPPEVFEQ:HQHSTGGWRLVNRLESWLETHALLE 180
 QY 193 KAUALVNGVPKASSPPHAYAVAILSDGQQRPEFQGPAGPAGSOGSSEGAOCLA 248
 DB 181 KAUALVNGVPKASSPPHAYAVAILSDGQQRPEFQGPAGPAGSOGSSEGAOCLA 236

RESULT 4
 ABB64965
 ID ABB64965 standard; Protein: 1398 AA.
 XX
 AC ABB64965;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 21699.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WC200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 LR WPI: 2001-656860/75.
 DR N-PSDB: ABL06352.
 XX

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
 PS Disclosure: SEQ ID NO 21699; 21pp - Sequence Listing: English.
 XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABH57737-ABH72072).
 CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1398 AA;

Query Match 41.9%; Score 551.5; DB 22; Length 1398;
 Best Local Similarity 48.4%; Pred. No. 6.5e-47;
 Matches 104; Conservative 40; Mismatches 64; Indels 7; Gaps 2;
 QY 3 LIAISPGINVKTFEDRMOLFSAIIGKPTPTPYEDGLYLFTQ:DNIIYPVPPHFCYLS 62
 DB 1132 MLKSSLPNGVVVYAYEDRMOLDSVMMVQPKRIFYNALFFDFDFQGRYPSPPVCHYIS 119;
 QY 63 QCSRLNPNLYDNKVCVSLGCTWAKGTEHWTSKSLLOVLSISQGLLVNRPYYNEAG 122

DB 1192 YCIDRLNPNLYEGGRVCVSLGCTWGRDNEWSPSCTMLQVLSIQGLLVDEPYNEAG 1251
 QY 123 FDSKRGLOEYNSRCYNEMALIRVQSM:TQIVRRPPEVFEQ:HQHSTGGWRLVNRLES 182
 DB 1252 YEKORTQLGNSRVRVYEMALIKIAQSTYVKQLINPLIFRNELIEHFKEFGTELYAMRK 1311
 QY 183 SWLETHAL-----LEKACALPNGVPKASSSPEPP 211
 DB 1312 AWSE-YSLAQKQNTIKTKDMPVRYKASCRLPEPP 1345

RESULT 5
 ABB62249
 ID ABB62249 standard; Protein: 4904 AA.
 XX
 AC ABB62249;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 13539.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WC200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 LR WPI: 2001-656860/75.
 DR N-PSDB: ABL06352.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
 PS Disclosure: SEQ ID NO 13539; 21pp - Sequence Listing: English.
 XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABH57737-ABH72072).
 CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 4904 AA;

Query Match 28.3%; Score 372.5; DB 22; Length 4904;
 Best Local Similarity 33.0%; Pred. No. 9.8e-28;
 Matches 98; Conservative 46; Mismatches 96; Indels 57; Gaps 10;
 QY 4 LATSIP-----EGINVKTFEDRMOLFSAIIGKPTPTPYEDGLYLFTQ:DNIIYPVPP--H 57
 DB 4572 LSTSLPUSFSVSVFVRCDTDRDLMKKVLIIGIADIPYANGCFEFDFFPPDPYQNPMLIN 4631
 QY 58 FCYLSQCSGRNPNLYDNKVCVSLGCTWAKGTEHWTSK-SLLQVLSISQGLLVNRP 116

ID 4632 LETTGRHSVREPNLYNDGKVCISVNTWICRPEEKWNAQSSFLQVLVSITQSLVPEP 4691
 QY 117 YYNEAGFSDRGLQKQYKNSCYNE---MALRVVQSMITQVRRPVPVEQERKQHFSTG 173
 DB 4692 YFNEPGFERSGRSPSCINSSREYNSTYQACVR--WAMEQIRSPSCQCFKUIWIKHFWLK 4749
 QY 174 GWELVNRKIESWLET-----HALLEKAKAL--PNG----- 200
 DB 4750 RHETCAQIEGWIPELCKKPCYTERASRTISNSMVLRRHYRHLREFLSK:KPPRGLFULD 4809
 QY 201 -----VFKASSPPFPAVAHSDSGOQPEGGVAP---GRASGSDSEGAAG 246
 DB 4810 PNPVALPPMDSVAVASLSA-ATNTVQAGPDALVTDIILLGENADECAAGDAEG 4865
 RESULT 6
 AAM80181
 ID AAM80181 standard; Protein: 527 AA.
 XX
 AC AAM80181;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 3827.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell; growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX Homo sapiens.
 CS
 XX
 PN W020015719C-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001: 2001WO-0504098.
 XX
 PR 03-FEB-2000: 2000US-0496914.
 PR 27-APR-2000: 2000US-0560875.
 PR 20-JUN-2000: 2000US-0598075.
 PR 19-JUL-2000: 2000US-0620325.
 PR 01-SEP-2000: 2000US-0654936.
 PR 15-SEP-2000: 2000US-0663561.
 PR 20-OCT-2000: 2000US-0693325.
 PR 30-NOV-2000: 2000US-0728422.
 XX (HYPF-) HYSEQ INC.
 XX
 PI Tang YI, Liu C, Drmanac RT, Asurdi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AC, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR N-PSDB: AAK5314.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20: Page 445: 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activator/inhibitor activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666

CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

SQ Sequence 527 AA;

Query Match 28.1%; Score 370; DB 22; Length 527;

Best Local Similarity 40.3%; Pred. No. 6.4e-29;

Matches 77; Conservative 36; Mismatches 68; Indels 10; Gaps 5;

QY 4 LATSIP-----EGIMYKTFEDAMDLFSLALIKQPTR:PYEDGILYFDIQ:PNLYPAVPHFC 59

DB 255 LSTSLPSSSSSVFRCDEERLDIMKVLITGPADIPYANGCEFDVFFQDYPSPP-IV 313

QY 60 YLSOCG---RLNPNIYDNGKVCVSLLTGKIGKSTERWISK-SSLQVLVLSIQGLIYNE 115

DB 314 NLETTGCHSVREPNLYNIGKVCISILNTHGHRPEKWNQPTSSSELQVLVSVSLIIIAE 373

QY 116 PYNAGFSDRGLQKQYKNSCYN-EMALIRVQVQSMITQVRRPVPVEQERKQHFSTG 174

DB 374 PYNEPGYERSRGTSGTQSSREYDNGIRQATVKWALEQINPNPSP:FEKVIHKKHYLKR 433

QY 175 WRLVNR:ESWL 185

DB 434 VEIMAQCEEWI 444

RESULT 7

AAB43174

ID AAB43174 standard; Protein: 1648 AA.

XX AAB43174;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2938 polypeptide sequence SEQ ID NO:5876.

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;

KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;

KW anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiac;

KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

KW hypotensive; dermatological; immunosuppressive; antinflammatory;

KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;

KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;

KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

KW cholesterol ester storage; systemic lupus erythematosus; infection;

KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

KW bone damage; cartilage damage; antinflammatory disease; coagulation;

KW thrombosis; contraceptive.

XX Homo sapiens.

CS

XX W0200058473-A2.

XX 05-OCT-2000.

PD

XX 31-MAR-2000: 2000WO-US08621.

XX 31-MAR-1999: 99US-0127607.

PR 02-APR-1999: 99US-0127636.

PR 05-APR-1999: 99US-0127728.

PR 30-MAR-2000: 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

PI WPI: 2000-602362/57.

DR N-PSDB: AAC77383.

XX Novel nucleic acids and peptides derived from open reading frame X,

XX useful for treating e.g. cancers, proliferative disorders,

PT

PI neurodegenerative disorders and cardiovascular disease -
 PS Claim 11: Page 5049-5053: 5507pp; English.
 XX AAC74446 to AAC77636 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiproliferative; antiparkinsonian; neutropenic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 1648 AA:
 Query Match 28.1%; Score 370; DB 21; Length 1648;
 Best Local Similarity 40.3%; Pred. No. 3.5e-28;
 Matches 77: Conservative 36; Mismatches 68; Indels 10; Gaps 5;
 QY 4 LATSIP---EGIMVKTFEDRMDFSAIKGTRIPYEDGLYFDIC:PNIIYPVPPHFC 59
 DB 1376 LSTSLPSSSSSVFVRCDEERLDIMKVLITGPAITPYANGCFEDYVFPQDPSSPP-LV 1434
 QY 60 YLSQCSG---RLNPNLYDNCKYCVSLCTGTWTKGTERWTSK-SSLLQVLISLOGLIYNE 115
 DB 1435 NLETTGSHSVRFNPNLYNDGKVCLSILNTWGRPEEKNPQTSFLOVLVSQSLILVAE 1494
 QY 116 PYNEAGFDSRGLQGYENSRNY-EMALIRVVQSMGLVRRPPEVFECE:RQHESTGG 174
 DB 1495 PYNEPGYENSRGTPSGTOSRNYDCNIRQATVWAMLEQIRNPSPCKEVIHKHEYLKR 1554
 QY 175 WFLVNRIFSWL 185
 DB 1555 VEIMAQCFEW 1565
 RESULT 8
 AAB9564
 ID AAB9564 standard: Protein; 1867 AA.
 XX AAB9564;
 AC AAB9564;
 XX 25-JUN-2003 (first entry)
 DE Human protein sequence SEQ ID NO:18204.
 XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 US EPI074617-A2.
 PN 07-FEB-2001.
 XX 28-JUL-2000; 2000JP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 21-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 PA Ola T, Isogai T, Mstikawa I, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-3:8749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8: SEQ ID 18204; 2537pp + CD ROM; English.
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAB03166 to AAB13628 and
 CC AAB13633 to AAB18742 represent human cDNA sequences; AAB92446 to
 CC AAB55893 represent human amino acid sequences; and AAB13629 to AAB13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 1867 AA:
 Query Match 28.1%; Score 370; DB 22; Length 1867;
 Best Local Similarity 40.3%; Pred. No. 4.2e-28;
 Matches 77: Conservative 36; Mismatches 68; Indels 10; Gaps 5;
 QY 4 LATSIP---EGIMVKTFEDRMDFSAIKGTRIPYEDGLYFDIC:PNIIYPVPPHFC 59
 DB 1595 LSTSLPSSSSSVFVRCDEERLDIMKVLITGPAITPYANGCFEDYVFPQDPSSPP-LV 1653
 QY 60 YLSQCSG---RLNPNLYDNCKYCVSLCTGTWTKGTERWTSK-SSLLQVLISLOGLIYNE 115
 DB 1654 NLETTGSHSVRFNPNLYNDGKVCLSILNTWGRPEEKNPQTSFLOVLVSQSLILVAE 1713
 QY 116 PYNEAGFDSRGLQGYENSRNY-EMALIRVVQSMGLVRRPPEVFECE:RQHESTGG 174
 DB 1714 PYNEPGYENSRGTPSGTOSRNYDCNIRQATVWAMLEQIRNPSPCKEVIHKHEYLKR 1773
 QY 175 WFLVNRIFSWL 185
 DB 1774 VEIMAQCFEW 1784
 RESULT 9
 AAB97833
 ID AAB97833 standard: Protein; 4829 AA.
 XX AAB97833;
 AC AAB97833;
 XX 09-AUG-2001 (first entry)
 DT Human apoptosis-inhibiting activity protein (IAP).
 XX Human; apoptosis-inhibiting activity; IAP; apoptosis inhibitor;

KW apoptosis.
 OS Homo sapiens.
 XX CP2001061484-A.
 XX 13-MAR-2001.
 XX
 PF 22 JUN-2000: 2000JP-C:87725.
 XX 23-JUN-1999: 99JP-0177200.
 XX
 PA (SANY J SANKYO CO LTD.
 EA (TSURU) TSURU T.
 XX WPI: 2001-294712/3.
 DR N-PSDB: AAB20191.
 XX
 PF Novel polynucleotide having apoptosis-inhibiting activity, used to
 PT treat diseases associated with abnormal apoptosis.
 XX
 PS Claim 2: Page 27-39; 44pp; Japanese.
 XX
 CC The present sequence represents a protein encoded by a polynucleotide
 CC comprising a 14490 nucleotide sequence (AAB20191) which has
 CC apoptosis-inhibiting activity. The polynucleotide can be used for the
 CC treatment and prevention of diseases caused by abnormal apoptosis.
 XX
 SQ Sequence 4829 AA;
 Query Match 28.1%; Score 370; DB 22; Length 4829;
 Rest Local Similarity 49.3%; Prod. No. 1-7e-27;
 Matches 77; Conservative 36; Mismatches 68; Indels 10; Gaps 5;
 QY 4 LATSLP-----ETIMVKTFEDRMDFSAIKGPTPTPYSDGILYFDIQINLYPAVPPHFC 59
 DB LSTSLPSSSSSVPRCOEERLDIMKVLITGPDITPYANGCFEDVYFPDYPSSPP-IV 4615
 QY 60 YLSCCSG----RLNPNLYNKGKVSILGIGWIKGIERWTSK-SSILQVLTISQGLILVNE 115
 DB LSTSLPSSSSSVPRCOEERLDIMKVLITGPDITPYANGCFEDVYFPDYPSSPP-IV 4615
 QY 116 PYNEAGTSDKGLQEGYENSRCYN-EMALIRVVSMTQLVRRPPEVFEQIRQHESTGG 174
 DB PFNPNPYSKSGIPSGTSSREYDNTGTCATVKKWAMLEQIRNPSPCFKEVHKHF 4735
 QY 175 WRVKNRTESWL 185
 DB VE:MAQCEEWI 4746
 RESULT 10
 AAB27993
 XX AAB27993 standard: Protein; 156 AA.
 XX AAB27993;
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE Human secreted protein BLAST search protein SEQ ID NO: 147.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 OS Homo sapiens.
 XX W0200055171 A1.
 XX
 PD 21-SEP-2000.
 XX

PF 09-MAR-2000: 2000WO-US96643.
 XX
 PR 12-MAR-1999: 99US-0124146.
 PR 23-NOV-1999: 99US-0167061.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI: 2000-638174/61.
 XX
 PI Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition.
 XX
 PS Disclosure; Page 418; 438pp; English.
 XX
 CC The invention relates to the isolation of genes AAC59049-C59098 encoding
 CC the human secreted proteins AAB27907-B27956. This sequence represents a
 CC peptide fragment homologous to the protein encoded by the gene given
 CC in the descriptor line. The sequence is a search result from a BLASTX
 CC homology search. The genes and proteins are useful for preventing,
 CC ameliorating or treating medical conditions, e.g. by protein or gene
 CC therapy. The genes are isolated from a range of human tissues disclosed
 CC in the specification. The nucleic acids, proteins, antibodies and
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such
 CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases
 CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such
 CC as viral, bacterial, fungal and parasitic infections.
 XX
 SQ Sequence 156 AA;
 Query Match 25.7%; Score 338; DB 21; Length 156;
 Rest Local Similarity 43.3%; Prod. No. 2e-26;
 Matches 68; Conservative 29; Mismatches 54; Indels 6; Gaps 4;
 QY 19 DRMDLSALIKGPTPTPYSDGILYFDIQINLYPAVPPHFCYLSQCSG---RLNPNLYDN 75
 DB 1 ERDIDKVLITGPDITPYANGCFEDVYFPDYPSSPP-IVNLETIGHSVRENPNLYND 59
 QY 76 GKVCVSLGTGCKTERWTSK-SSILQVLTISQGLILVNEPYNEAGTSDKGLQEGYFE 134
 DB 60 GKVCLSILMTWGRPEEKWNPOTSSFLQVLSVQSILVAEPYNEPGRYSKGFESGIQ 119
 QY 135 NSRCYN-EMALIRVVSMTQLVRRPPEVFEQIRQHF 170
 DB 120 SSREYDNTGTCATVKKWAMLEQIRNPSPCFKEVHKHF 156
 RESULT 11
 AAB94855
 XX AAB94855 standard: Protein; 246 AA.
 XX AAB94855;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:16042.
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 XX EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000: 2000EP-C1:6126.

XX 29-JUL-1999; 95JP-0248036.
 PR 27-AUG-1999; 95JP-0300253.
 PR 11-JAN-2000; 2000JP-018776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELIX-) HELIX RES INST.
 XX Ota T, Isoqai T, Nishikawa I, Hayashi K, Saito K, Yamamoto J;
 PI Shii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WP1: 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PI full-length cDNAs defined in the specification, and for the detection
 PI and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs
 XX Claim 8: SEQ ID 16042: 2337bp - CD ROM; English.
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH5893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX Sequence 246 AA:
 SQ
 Query Match 18.9%; Score 248; Dh 22; Length 246;
 Best Local Similarity 43.7%; Pred. No. 6.7e-17;
 Matches 59; Conservative 16; Mismatches 50; Indels 10; Gaps 5;
 QY 9 PEGIMVKTFEDRMDFLSALIKGPTPTPEYDGLYLFDTQLNPIYPAVPPHFCYLSQCSG-- 66
 DB 8 PPGMFWPDPVDMTKIHAILITGPDPIPEGGGFLFVRCPDPYTHPPRVKLMITGNTV 67
 QY 67 RLNPILYDNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLISIQGLILVNEPYNAGFDS 126
 DB 68 RNPFPYRNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLISIQGLILVNEPYNAGFDS 126
 QY 127 RGLQEGYENSCYNE 141
 DB 224 R--HPG--DSKNYNE 134
 Query Match 18.9%; Score 248; Dh 22; Length 295;
 Best Local Similarity 43.7%; Pred. No. 8.8e-17;
 Matches 59; Conservative 16; Mismatches 50; Indels 10; Gaps 5;
 QY 9 PEGIMVKTFEDRMDFLSALIKGPTPTPEYDGLYLFDTQLNPIYPAVPPHFCYLSQCSG-- 66
 DB 129 PPGMFWPDPVDMTKIHAILITGPDPIPEGGGFLFVRCPDPYTHPPRVKLMITGNTV 188
 QY 67 RLNPILYDNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLISIQGLILVNEPYNAGFDS 126
 DB 189 RNPFPYRNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLISIQGLILVNEPYNAGFDS 126
 QY 127 RGLQEGYENSCYNE 141
 DB 245 R--HPG--DSKNYNE 255

Human; molecules for disease detection and treatment; mddt;
 KW Antiarteriosclerotic; hepatotropic; antiprosoratic; cytosclastic;
 KW immunosuppressive; antidiabetic; antihistaminic; neutrophilic;
 KW osteopathic; antiarthritic; cell proliferative disorder;
 KW arteriosclerosis; cirrhosis; psoriasis; cancer; adenocarcinoma;
 KW leukaemia; breast cancer; autoimmune disorder; AIDS;
 KW acquired immunodeficiency syndrome; Addison's disease;
 KW diabetes mellitus; asthma; multiple sclerosis; osteoarthritis.
 XX Homo sapiens.
 OS WO200162922-A2.
 XX 30-AUG-2001.
 XX 21-FEB-2001; 2001WO-US05896.
 XX 24-FEB-2000; 2000US-0185213.
 PR 16-MAY-2000; 2000US-0205232.
 PR 17-MAY-2000; 2000US-0205285.
 PR 17-MAY-2000; 2000US-0205286.
 PR 17-MAY-2000; 2000US-0205287.
 PR 17-MAY-2000; 2000US-0205323.
 PR 17-MAY-2000; 2000US-0205324.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Panzer SR, Spiro PA, Barville SC, Shah P, Chalup MS, Chang SC;
 PI Chen A, D'Sa SA, Ansley S, Dahl CR, Dam TC, Daniels SE;
 PI Dufour GE, Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL;
 PI Liu TF, Roseberry AM, Rosen BL, Russo FD, Stockreiner TK, Daffo A;
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Braucher SR, Chen W;
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;
 XX WP1: 2001-570631/54.
 DR N-PSDB; AAS42509.
 XX New disease detection and treatment molecule polynucleotides and
 PT polypeptides, useful for diagnosis and treatment of arteriosclerosis,
 PT cirrhosis, psoriasis, cancer, autoimmune disorders, diabetes mellitus,
 PT asthma and multiple sclerosis
 XX Claim 27; Page 169; 183pp; English.
 XX The invention relates to novel human molecules for disease
 CC detection and treatment (mddt proteins) and the polynucleotides encoding
 CC them. The mddt polynucleotides and polypeptides are useful for diagnostic
 CC and therapeutic purposes e.g. to diagnose and treat cell proliferative
 CC disorders (e.g. arteriosclerosis, cirrhosis and psoriasis) cancers (e.g.
 CC adenocarcinoma, leukaemia and breast cancer) autoimmune disorders
 CC (e.g. acquired immunodeficiency syndrome (AIDS) and Addison's disease)
 CC diabetes mellitus, asthma, multiple sclerosis, osteoarthritis, and many
 CC more diseases given in the specification. The present sequence
 CC represents an mddt protein of the invention.
 XX Sequence 295 AA:
 Query Match 18.9%; Score 248; Dh 22; Length 295;
 Best Local Similarity 43.7%; Pred. No. 8.8e-17;
 Matches 59; Conservative 16; Mismatches 50; Indels 10; Gaps 5;
 QY 9 PEGIMVKTFEDRMDFLSALIKGPTPTPEYDGLYLFDTQLNPIYPAVPPHFCYLSQCSG-- 66
 DB 129 PPGMFWPDPVDMTKIHAILITGPDPIPEGGGFLFVRCPDPYTHPPRVKLMITGNTV 188
 QY 67 RLNPILYDNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLISIQGLILVNEPYNAGFDS 126
 DB 189 RNPFPYRNGKVCVSLIGTWIGKGTERTWTSKSSLLQVLISIQGLILVNEPYNAGFDS 126
 QY 127 RGLQEGYENSCYNE 141
 DB 245 R--HPG--DSKNYNE 255

RESULT 13

AAB94781
ID AAB94781 standard; Protein; 236 AA.

XX AAB94781

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:15879.

XX Human: primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EPI074617-A2.

XX 07-FEB 2001.

XX 28-JUL-2000; 2000EP-011E126.

XX 29-JUL-1999; 99JP-024E336.

XX 27-AGO-1999; 99JP-030E253.

XX 11-JAN-2000; 2000JP-018776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Iwatai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Naqai K, Otsuki T;

XX WPI; 2001-316749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

XX Claim 8: SEQ ID 15879; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in:
CC the specification. The primer sets can be used in antisense therapy and
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAB03166 to AAB13628 and
CC AAB13633 to AAB18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAB13629 to AAB13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 236 AA:

XX Query Match 18.6%; Score 244; DB 22; Length 236;
XX Best Local Similarity 44.9%; Pred. No. 1.6e-16;
XX Matches 57; Conservative 17; Mismatches 4; Indels 12; Gaps 6:

XX 19 DRMLC--FSALNGPRTPTDYLIDQLINPVAVPHFCYLSQSG--R--NPNLXD 74

DB 6 STVMTKIHAIITGPDPTPGSGFLFVRCPPDPYIHPVRVKLMTGNNVRENPNFYR 65

OY 75 NGKVCVSLLS--TWIGKGTERTWTSKSSLLQVLISIQGILVNEPYNENAGFLSDRGLOEYVE 134

DB 66 NGKVCISILGWTGPA---WSPAQSISSVLSIQSLMTEN-PYHNKPGFEQHR--HPG-- 117

OY 135 NSRCYNE 141

DB 118 DSKYNE 124

RESULT 14

AAB27994

ID AAB27994 standard; Protein; 152 AA.

XX AAB27994;

XX 02-FEB-2001 (first entry)

XX Human secreted protein BLAST search protein SEQ ID NO: 148.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

XX WO2000055171-A1.

XX 21-SEP-2000.

XX 09-MAR-2000; 2000WO-US060043.

XX 12-MAR-1999; 99US-0124146.

XX 23-NOV-1999; 99US-0167061.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-638174/61.

XX Isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition -

XX Disclosure; page 419-420; 438pp; English.

XX The invention relates to the isolation of genes AAC59049-C59098 encoding
CC the human secreted proteins AAB27907-B27956. This sequence represents a
CC fragment of the protein encoded by the gene given in the descriptor
CC line. The sequence is used as a query sequence for doing BLASTX searches
CC to determine homologous sequence to the protein. The genes and proteins
CC are useful for preventing, ameliorating or treating medical conditions
CC e.g. by protein or gene therapy. The genes are isolated from a range of
CC human tissues disclosed in the specification. The nucleic acids,
CC proteins, antibodies and (antagonists are useful in the diagnosis,
CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,
CC and other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and
CC (f) infectious diseases such as viral, bacterial, fungal and parasitic
CC infections.

XX Sequence 152 AA:

XX Query Match 15.4%; Score 203; DB 21; Length 152;
XX Best Local Similarity 38.6%; Pred. No. 1.4e-12;
XX Matches 49; Conservative 16; Mismatches 50; Indels 12; Gaps 6:

CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX Sequence 283 AA;
 SQ
 Query Match 12.9%; Score 169; DB 21; Length 283;
 Best Local Similarity 24.5%; Pred. No. 1.1e-08;
 Matches 60; Conservative 36; Mismatches 73; Indels 76; Gaps 10;
 QY 4 LATSIPGIMVKTPEDEMDLFSALIKGPTRTYEEGLYLFDIQIPN:YPAVPHPFCYLSQ 63
 DB 23 LTADPPGIRKVPNEEDLTDLQVTEGPEG:PYAGGLFRMKLLGKDFPASPCKGYFTK 82
 QY 64 CSGRLNPNLYDNGKVCVSLGWTGKGTERTWTSKSSLLQVLISIQGLILVNEPYNEAGF 123
 DB 83 L----PHPNVGANGECVNVL-----KRDWTAELGIRHVLITIKCLLHPNP-----ESAL 129
 QY 124 QSDRG--LQEGYNSKCYNEMALIRVQSMITQIVRRPPRVFQEIHQHFSIGGWRLVNRI 181
 DB 130 NEEAGRLLENYEE---YAARL-----LTEI-----HGGAGG----- 160
 QY 182 ESWLETHALLEKQAQ:PNGVFPKASSPPPPAVAEISDSGQCHPEDGPGAPGASQGSUSE 241
 DB 161 -----PSGRABAGRALASGTEASSTDPGAP-----CGP----- 188
 QY 242 GGAQG 246
 DB 189 GGAQS 193
 RESULT 17
 AAB43534
 ID AAB43534 standard; Protein; 283 AA.
 AC AAB43534;
 XX
 XX 08-FEB-2002 (first entry)
 DE Human cancer associated protein sequence SEQ ID NO:979.
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; antineoplastic; immunomodulator;
 KW antidiabetic; antihistaminic; antihypertensive; antithrombotic; antiviral;
 KW antiinflammatory; antihypertensive; antihypertensive; antihypertensive;
 KW dermatological; neoplastic; thrombolytic; coagulant; neotropic;
 KW vasotropic; antipsoriasis; antianqurogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX Homo sapiens.
 CS
 XX W0200055550-A1.
 PN
 XX 21-SEP-2000.
 PF
 XX 08-MAR-2000; 2000WO-US05842.
 PR
 XX 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 XX
 XX WPI: 2000-587533/55.
 DR
 XX N-PSDB; AAC77743.
 XX
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 PT
 XX
 XX Claim 11: Page 1551: 1552: 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC includes: cytostatic; proliferative; antineoplastic; immunomodulator;
 CC antidiabetic; antihistaminic; antihypertensive; antithrombotic;
 CC antiinflammatory; antihypertensive; antihypertensive; antihypertensive;
 CC dermatological; neoplastic; thrombolytic; coagulant; neotropic;
 CC vasotropic; antipsoriasis; antianqurogenic; gene therapy; inflammation;
 CC immune disorder; haematopoietic cell disorder; autoimmune disorder;
 CC allergic reaction; graft versus host disease; organ rejection;
 CC haemostatic; thrombolytic; cardiovascular disorder; infection;
 CC neurological disease; drug screening.

CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX Sequence 283 AA;
 SQ
 Query Match 12.9%; Score 169; DB 21; Length 283;
 Best Local Similarity 24.5%; Pred. No. 1.1e-08;
 Matches 60; Conservative 36; Mismatches 73; Indels 76; Gaps 10;
 QY 4 LATSIPGIMVKTPEDEMDLFSALIKGPTRTYEEGLYLFDIQIPN:YPAVPHPFCYLSQ 63
 DB 84 LTADPPGIRKVPNEEDLTDLQVTEGPEG:PYAGGLFRMKLLGKDFPASPCKGYFTK 143
 QY 64 CSGRLNPNLYDNGKVCVSLGWTGKGTERTWTSKSSLLQVLISIQGLILVNEPYNEAGF 123
 DB 144 L----PHPNVGANGECVNVL-----KRDWTAELGIRHVLITIKCLLHPNP-----ESAL 190
 QY 124 QSDRG--LQEGYNSKCYNEMALIRVQSMITQIVRRPPRVFQEIHQHFSIGGWRLVNRI 181
 DB 191 NEEAGRLLENYEE---YAARL-----LTEI-----HGGAGG----- 221
 QY 182 ESWLETHALLEKQAQ:PNGVFPKASSPPPPAVAEISDSGQCHPEDGPGAPGASQGSUSE 241
 DB 222 -----PSGRABAGRALASGTEASSTDPGAP-----CGP----- 249
 QY 242 GGAQG 246
 DB 250 GGAQS 254
 RESULT 18
 AAG75556
 ID AAG75556 standard; Protein; 285 AA.
 AC AAG75556;
 XX
 XX 03-SEP-2001 (first entry)
 DE Human colon cancer antigen protein SEQ ID NO:6320.
 XX
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 17.
 XX
 XX Homo sapiens.
 OS
 XX W0200122920-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 28-SEP-2000; 2000WO-US26524.
 PF
 XX 29-SEP-1999; 99US-C157137.
 PR
 XX 03-NOV-1999; 99US-0163280.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX WPI: 2001-235357/24.
 DR
 XX N-PSDB; AAB34361.
 XX
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides.

PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX Claim 11: Page 7778-7779; 9803pp; English.
 PS
 XX
 CC AAB32943 to AAB37135 and AAB373514 to AAB37788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAB37196 to AAB37204
 CC and AAB37789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1627 to 1052, 7921 and 7922.
 XX

Sequence 285 AA:

Query Match 12.9% Score 169; DB 22; Length 285;
 Best Local Similarity 24.5%; Pred. No. 1.1e-08;
 Matches 60; Conservative 36; Mismatches 73; Indels 76; Gaps 10;
 QY 4 LATSIPGIMVKTFEDRMDFLSALIKGTRTPYEDGLYLFIDQ:PNVAVPPHFCYLSQ 63
 DB 86 LTADPPDGKIVFPNEEDLTLQVITGEPGTPYAGGLFPMKLLIGKDFPASPFGYFTK 145
 QY 64 CSRLPNPNLYDNGKVCVSLLTGWTGKTERWTSKSLLOVLISQGLILVNEPYNEAGF 123
 DB 146 I---FHPNVGANCEICVNVL-----KRWIAELGIRHVLLIKCLLIHPNP---ESAL 192
 QY 124 ESDRG---LQEGVENSRCYNEMALIRVVQSMTQVLRPPPEVFEQIRHSTGGWRLVNR 181
 DB 193 NEAGRLLENTEE---YNAARL-----LTEI-----HGGAGG----- 223
 QY 162 ESWLETHALCEKAQALPNCVTKASSPEPPAVAEILSDSGQHPEDGCPAFGRASQGSUSE 241
 DB 224 ---PSGRAEGRALASSTASSTIDGAP-----GGP----- 251
 QY 242 GGAAG 246
 DB 252 GGAAG 256

RESULT 19

ABP41918
 ID ABP41918 standard: Protein; 285 AA.

XX ABP41918:

XX
 XX
 DT 22-ARG-2002 (first entry)

XX Human ovarian antigen HGO0F6; SEQ ID NO:3050.

KW Human ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 17.
 XX Homo sapiens.

XX
 PN
 XX
 PD
 XX
 PF
 XX
 PR
 XX
 PA
 XX
 PI
 XX
 DR
 XX
 DR
 XX
 PT
 XX
 PT
 XX
 PT
 XX
 PS
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054 -
 CC ABP43228) and to cDNAs encoding them (ABP4111-ABP436305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 285 AA:

Query Match 12.9% Score 169; DB 23; Length 285;
 Best Local Similarity 24.5%; Pred. No. 1.1e-08;
 Matches 60; Conservative 36; Mismatches 73; Indels 76; Gaps 10;
 QY 4 LATSIPGIMVKTFEDRMDFLSALIKGTRTPYEDGLYLFIDQ:PNVAVPPHFCYLSQ 63
 DB 86 LTADPPDGKIVFPNEEDLTLQVITGEPGTPYAGGLFPMKLLIGKDFPASPFGYFTK 145
 QY 64 CSRLPNPNLYDNGKVCVSLLTGWTGKTERWTSKSLLOVLISQGLILVNEPYNEAGF 123
 DB 146 I---FHPNVGANCEICVNVL-----KRWIAELGIRHVLLIKCLLIHPNP---ESAL 192
 QY 124 ESDRG---LQEGVENSRCYNEMALIRVVQSMTQVLRPPPEVFEQIRHSTGGWRLVNR 181
 DB 193 NEAGRLLENTEE---YNAARL-----LTEI-----HGGAGG----- 223
 QY 182 ESWLETHALCEKAQALPNCVTKASSPEPPAVAEILSDSGQHPEDGCPAFGRASQGSUSE 241
 DB 224 ---PSGRAEGRALASSTASSTIDGAP-----GGP----- 251
 QY 242 GGAAG 246
 DB 252 GGAAG 256

Db 224 -----PSGRAPAGRALASGTEASSTDPAP -----GCP----- 251
Qy 242 GAGOG 246
Db 252 GAGG 256

RESULT 20
ID: AAG10072 standard; Protein: 153 AA.
XX AAG10072;
AC AAG10072;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 8249.
XX protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX Arabidopsis thaliana.
XX EP1033425-A2.
XX 66-SEP-2003.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130377.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130591.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 20-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142350.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144884.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 22-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 23-JUL-1999; 99US-0145224.
XX 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
XX 27-JUL-1999; 99US-0145918.
XX 27-JUL-1999; 99US-0145919.
XX 28-JUL-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146386.
XX 02-AUG-1999; 99US-0146388.
XX 02-AUG-1999; 99US-0146389.
XX 03-AUG-1999; 99US-0147038.
XX 04-AUG-1999; 99US-0147204.
XX 04-AUG-1999; 99US-0147302.
XX 05-AUG-1999; 99US-0147192.
XX 06-AUG-1999; 99US-0147260.
XX 06-AUG-1999; 99US-0147303.
XX 06-AUG-1999; 99US-0147416.
XX 09-AUG-1999; 99US-0147453.
XX 09-AUG-1999; 99US-0147535.
XX 10-AUG-1999; 99US-0148171.
XX 11-AUG-1999; 99US-0148319.
XX 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148564.
 PR 16-AUG-1999; 99US-0143368.
 PR 17-AUG-1999; 99US-0143175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0143722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151085.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154058.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155133.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158363.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159337.
 PR 14-OCT-1999; 99US-0159367.
 PR 18-OCT-1999; 99US-0159384.
 PR 21-OCT-1999; 99US-0160747.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160380.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161494.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161320.
 PR 28-OCT-1999; 99US-0161392.
 PR 28-OCT-1999; 99US-0161393.
 PR 29-OCT-1999; 99US-0162142.

Query Match 12.6%; Score 166; DB 21; Length 153;
 Best Local Similarity 28.4%; Pred. No. 8.6e-09;
 Matches 43; Conservative 27; Mismatches 62; Indels 20; Gaps 4;
 QY 3 LLAISLPEIGIMVKIFEDMDIFSAICKGTRKYFEDXGILYFDICJPMIVPAYERPHCYLS 62
 DB 17 LLSKAP-CISASPSDMMRYFNMVLI:GTSUYEGGVKLELELPDEYPMAPKVRFLI 75
 QY 63 QCGRLNPLNDKXCVSLIGTWIGKGTERTWSKSLLOVLISIQGLILVNEPYTNEAG 122

Db 76 KI--YHPNICKLGRICLDIL-----KRWSPALQIRTVLLSIGILASNP----- 119
 QY 123 FSDRGIAEGYNSKYNEMALIRVVQSMIOL 154
 Db 120 ---DDPLSENIAKHWSNEAEAVDTAKEWIRL 148
 RESULT 2:
 AAG48920
 ID AAG48920 standard; Protein: 153 AA.
 XX
 AC AAG48920;
 XX
 DI 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SHQ ID NO: 61831.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128214.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.


```

RESULT 24
AAY23157
ID AAY23157 standard; Protein: 197 AA.
XX AC AAY23157;
XX DT 26-AUG-1999 (first entry)
XX DE Human ubiquitin-like conjugating protein (UBCLE).
XX KW Human; ubiquitin-like conjugating protein; UBCL; cancer;
XX KW developmental disorder; immune disorder; neuronal disorder;
XX KW neuronal protein; Alzheimer's disease; Down's Syndrome;
XX KW Parkinson's disease; Cushing's disease; renal tubular acidosis;
XX KW Duchenne's muscular dystrophy; Addison's disease;
XX KW ankylosing spondylitis; Hashimoto's thyroiditis.
XX OS Homo sapiens.
XX PN WC9331252-A1.
XX PP 24-JUN-1999.
XX PF 12-DEC-1998; 98WO-0525564.
XX PR 12-DEC-1997; 97CS-0963289.
XX PA (TNCY-) INCYTE PHARM INC.
XX PI Corley NC, Hillman JL, Shah P;
XX DR WPI: 1999-495347/34.
XX DX N-PSDB; AAX81676.
XX PT New human protein useful for diagnosing and treating cancers.
XX PS neuronal disorders, immune disorders and developmental disorders
XX PS Claim 1; Fig 1A-C; 6pp; English.
XX CC The present sequence represents human ubiquitin-like conjugating protein;
XX CC (UBCLE). The protein may be administered to human patients or mammals
XX CC such as monkeys or farm animals to treat or prevent cancer, a
XX CC developmental disorder, an immune disorder or a neuronal disorder
XX CC resulting from an abnormality in the processing of neuronal proteins.
XX CC Examples of cancers which may be treated are adrenal, bladder,
XX CC bone, brain ovarian, prostate and uterus. Examples of neuronal disorders
XX CC which may be treated are Alzheimer's disease, Down's Syndrome and
XX CC Parkinson's disease. Examples of developmental disorders which may be
XX CC treated are Cushing's syndrome, renal tubular acidosis and Duchenne's
XX CC muscular dystrophy. Examples of immune disorders which may be
XX CC treated are Addison's disease, ankylosing spondylitis and Hashimoto's
XX CC thyroiditis.
XX SQ Sequence 197 AA;
Query Match 12.5%; Score 164; DB 20; Length 197;
Best Local Similarity 28.3%; Pred. No. 2e-08;
Matches 49; Conservative 33; Mismatches 75; Indels 16; Gaps 6;
QY 3 LLATSLPEGTWVKIFEDRMDLFSALINGKPTKTYEDGLYFDIQLPNIYPAVPHFCYLS 62
DB 13 MIATPPPGTICWQKQDMDLRAVLGGANTPYKGVKLEVIIPERYPEPPQIRFLT 72
QY 63 QCSRLPNLYDNGKGVSLGTLGKGTERTKTSKSLLOVLISIQCLLYNPNYNEAG 122
DB 73 PT --HPNIDNAGR:CLDVA, KLPTPGA---WRPSINATVLTISIQ--LWSEP----- 118
QY 123 FDSRGLQGYNSRCYNEMALIRIVVQSMTQL-VRRPPEVFEQETROHSTGS 174
DB 115 -NPDPIPLADISSEPKYKPKAPLANAKQWTEKHAKOKRADEEFLMDLNLPEAG 170

```

```

RESULT 25
AAM39669
ID AAM39669 standard; Protein: 197 AA.
XX AC AAM39669;
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 2814.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukaemia.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PP 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-058042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX DR WPI: 2001-442253/47.
XX DX N-PSDB; AAI58825.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PS such as central nervous system injuries -
XX PS Example 4; SEQ ID NO 2814; 10078pp; English.
XX CC The invention relates to human nucleic acids (AA157798-AA161369) and
XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: Immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.
XX SQ Sequence 197 AA;
Query Match 12.5%; Score 164; DB 22; Length 197;
Best Local Similarity 28.3%; Pred. No. 2e-08;
Matches 49; Conservative 33; Mismatches 75; Indels 16; Gaps 6;
QY 3 LLATSLPEGTWVKIFEDRMDLFSALINGKPTKTYEDGLYFDIQLPNIYPAVPHFCYLS 62

```


DB 119 - NPDDPILMADISSPKYNKAF..KNARQWTEKHARQKQKAEFEFMDLNUPEAG 170
 RESULT 2H
 AAU23245
 ID AAU23245 standard; Protein; 207 AA.
 AC AAU23245;
 XX 17-DEC-2001 (first entry)
 XX Novel human enzyme polypeptide #33.
 XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW lique; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW nephrotropic; anticoagulant.
 XX Homo sapiens.
 XX WC20010530; A2.
 XX 02-AUG-2001.
 XX 17-JAN-2001; 2001W0-0501239.
 XX 31-JAN-2000; 2000US-0179065.
 XX 04-FEB-2000; 2000US-0180628.
 XX 24-FEB-2000; 2000US-0184564.
 XX 02-MAR-2000; 2000US-0186350.
 XX 16-MAR-2000; 2000US-0189874.
 XX 17-MAR-2000; 2000US-0190076.
 XX 18-APR-2000; 2000US-0198123.
 XX 19-MAY-2000; 2000US-0205515.
 XX 07-JUN-2000; 2000US-0209467.
 XX 28-JUN-2000; 2000US-0214886.
 XX 30-JUN-2000; 2000US-0215135.
 XX 07-JUL-2000; 2000US-0216647.
 XX 07-JUL-2000; 2000US-0216882.
 XX 11-JUL-2000; 2000US-0217487.
 XX 11-JUL-2000; 2000US-0217496.
 XX 14-JUL-2000; 2000US-0218230.
 XX 26-JUL-2000; 2000US-0220953.
 XX 26-JUL-2000; 2000US-0220956.
 XX 14-AUG-2000; 2000US-0224518.
 XX 14-AUG-2000; 2000US-0224519.
 XX 14-AUG-2000; 2000US-0225213.
 XX 14-AUG-2000; 2000US-0225214.
 XX 14-AUG-2000; 2000US-0225266.
 XX 14-AUG-2000; 2000US-0225267.
 XX 14-AUG-2000; 2000US-0225268.
 XX 14-AUG-2000; 2000US-0225270.
 XX 14-AUG-2000; 2000US-0225447.
 XX 14-AUG-2000; 2000US-0225757.
 XX 14-AUG-2000; 2000US-0225758.
 XX 14-AUG-2000; 2000US-0225759.
 XX 18-AUG-2000; 2000US-0226279.
 XX 22-AUG-2000; 2000US-0226681.
 XX 22-AUG-2000; 2000US-0226868.
 XX 22-AUG-2000; 2000US-0227182.
 XX 23-AUG-2000; 2000US-0227009.
 XX 30-AUG-2000; 2000US-0228924.
 XX 01-SEP-2000; 2000US-0229287.
 XX 01-SEP-2000; 2000US-0229343.
 XX 01-SEP-2000; 2000US-0229344.
 XX 01-SEP-2000; 2000US-0229345.
 XX 05-SEP-2000; 2000US-0229509.
 XX 05-SEP-2000; 2000US-0229513.
 XX 06-SEP-2000; 2000US-0230437.
 XX 06-SEP-2000; 2000US-0230438.
 XX 08-SEP-2000; 2000US-0231242.
 XX 08-SEP-2000; 2000US-0231243.
 XX 08-SEP-2000; 2000US-0231244.
 XX 08-SEP-2000; 2000US-0231413.
 XX 08-SEP-2000; 2000US-0231414.
 XX 08-SEP-2000; 2000US-0232080.
 XX 12-SEP-2000; 2000US-0232081.
 XX 14-SEP-2000; 2000US-0231968.
 XX 14-SEP-2000; 2000US-0232397.
 XX 14-SEP-2000; 2000US-0232398.
 XX 14-SEP-2000; 2000US-0232399.
 XX 14-SEP-2000; 2000US-0232400.
 XX 14-SEP-2000; 2000US-0232401.
 XX 14-SEP-2000; 2000US-0233063.
 XX 14-SEP-2000; 2000US-0233064.
 XX 14-SEP-2000; 2000US-0233065.
 XX 21-SEP-2000; 2000US-0234223.
 XX 21-SEP-2000; 2000US-0234274.
 XX 25-SEP-2000; 2000US-0234997.
 XX 26-SEP-2000; 2000US-0234998.
 XX 27-SEP-2000; 2000US-0235484.
 XX 27-SEP-2000; 2000US-0235834.
 XX 29-SEP-2000; 2000US-0235836.
 XX 29-SEP-2000; 2000US-0236327.
 XX 29-SEP-2000; 2000US-0236367.
 XX 29-SEP-2000; 2000US-0236368.
 XX 29-SEP-2000; 2000US-0236369.
 XX 29-SEP-2000; 2000US-0236370.
 XX 02-OCT-2000; 2000US-0236802.
 XX 02-OCT-2000; 2000US-0237037.
 XX 02-OCT-2000; 2000US-0237038.
 XX 02-OCT-2000; 2000US-0237039.
 XX 02-OCT-2000; 2000US-0237040.
 XX 13-OCT-2000; 2000US-0239935.
 XX 13-OCT-2000; 2000US-0239937.
 XX 20-OCT-2000; 2000US-0240960.
 XX 20-OCT-2000; 2000US-0241221.
 XX 20-OCT-2000; 2000US-0241785.
 XX 20-OCT-2000; 2000US-0241786.
 XX 20-OCT-2000; 2000US-0241787.
 XX 20-OCT-2000; 2000US-0241808.
 XX 20-OCT-2000; 2000US-0241809.
 XX 20-OCT-2000; 2000US-0241826.
 XX 01-NOV-2000; 2000US-0244617.
 XX 08-NOV-2000; 2000US-0244674.
 XX 08-NOV-2000; 2000US-0244675.
 XX 08-NOV-2000; 2000US-0246476.
 XX 08-NOV-2000; 2000US-0246477.
 XX 08-NOV-2000; 2000US-0246478.
 XX 08-NOV-2000; 2000US-0246523.
 XX 08-NOV-2000; 2000US-0246524.
 XX 08-NOV-2000; 2000US-0246525.
 XX 08-NOV-2000; 2000US-0246526.
 XX 08-NOV-2000; 2000US-0246527.
 XX 08-NOV-2000; 2000US-0246528.
 XX 08-NOV-2000; 2000US-0246532.
 XX 08-NOV-2000; 2000US-0246609.
 XX 08-NOV-2000; 2000US-0246610.
 XX 08-NOV-2000; 2000US-0246611.
 XX 08-NOV-2000; 2000US-0246613.
 XX 17-NOV-2000; 2000US-0249207.
 XX 17-NOV-2000; 2000US-0249208.
 XX 17-NOV-2000; 2000US-0249209.
 XX 17-NOV-2000; 2000US-0249210.
 XX 17-NOV-2000; 2000US-0249211.
 XX 17-NOV-2000; 2000US-0249212.
 XX 17-NOV-2000; 2000US-0249213.
 XX 17-NOV-2000; 2000US-0249214.
 XX 17-NOV-2000; 2000US-0249215.
 XX 17-NOV-2000; 2000US-0249216.
 XX 17-NOV-2000; 2000US-0249217.
 XX 17-NOV-2000; 2000US-0249218.
 XX 17-NOV-2000; 2000US-0249244.
 XX 17-NOV-2000; 2000US-0249245.

```

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249303.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251033.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251356.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HOMA-) HUMAN GENOME SCI INC.
PA
PI Rosen CA, Harash SG, Reben SM;
XX
XX
DR WPI: 2001-465566/50.
DR N-PSDB: AAS41115.
XX
XX
XX Novel polypeptides and polynucleotides useful for diagnosing,
PI preventing, treating neural, immune system, muscular, reproductive,
PI pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PI diseases.
XX
XX Claim 11: SEQ ID NO 124i: 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerase or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC blood-related disorders (e.g. haemophilia), reproductive disorders
CC (e.g. infertility) and infectious disorders (e.g. influenza). The
CC polynucleotides of the invention can also be used in gene therapy.
CC AA022915-AA023814 represent the novel human enzyme polypeptides of the
CC invention.
XX
XX Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Query Match 12.5%; Score 164; DB 22; Length 207;
Ref Local Similarity 26.3%; Pred. No. 2.2e-08;
Matches 49; Conservative 33; Mismatches 75; Indels 16; Gaps 6;

QY 3 ILATSLPEGIWVKIFEDRMILFSAITKGTPTPYEDSLYFLDQLPNLYPAVPPHFCYLS 62
Db :||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23 MLATEPPPGITCQCKDQMDLRAOILGGANTPEKGVKLEVIIPERYPEPQIRFLI 82
QY 63 QCSGMLPNIYDNGKVCVSLIG:WICKGIERTWTSKSLQVLISIOGLIYNENLYPNFAG 122
Db :||| :||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 PI-- YIENICSAKICLQVLI-KLPKKA -WRPSLIATVLTISQ--LIMSEP----- 128

QY 123 FDSURGLQEGYNSRCYNEMALIRVVQSWTQ:-VRRPEVFEOEIRQHFSTGG 174
Db :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 -NEDPIIMADISSEFFRYNKPAPLKNAROWIEKHARQKQKADEEMLDNLPFAG 180

RESULT 29

```

```

AAM41455
ID AAM41455 standard; Protein; 207 AA.
XX
AC AAM41455;
XX
DI 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6386.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
XX W0200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000GS-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0633450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HVSFO INC.
XX
XX Tanq YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
XX N-PSDB: AAT60611.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries.
XX
XX Example 2; SEQ ID NO 6386; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 207 AA:

```

```

Query Match 12.5%; Score 164; DB 22; Length 207;
Best Local Similarity 26.3%; Pred. No. 2.2e-08;
Matches 49; Conservative 33; Mismatches 75; Indels 16; Gaps 6;

QY 3 ILATSLPEGIWVKIFEDRMILFSAITKGTPTPYEDSLYFLDQLPNLYPAVPPHFCYLS 62
Db :||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23 MLATEPPPGITCQCKDQMDLRAOILGGANTPEKGVKLEVIIPERYPEPQIRFLI 82

RESULT 29

```


PR 06-MAY-1999; 990S-0132457.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134373.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135620.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139452.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140354.
PR 28-JUN-1999; 990S-0140895.
PR 29-JUN-1999; 990S-0140823.
PR 30-JUN-1999; 990S-0140951.
PR 01-JUL-1999; 990S-0141287.
PR 03-JUL-1999; 990S-0141842.
PR 03-JUL-1999; 990S-0142154.
PR 03-JUL-1999; 990S-0142085.
PR 06-JUL-1999; 990S-0142350.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144425.
PR 19-JUL-1999; 990S-0144433.
PR 19-JUL-1999; 990S-0144432.
PR 19-JUL-1999; 990S-0144433.
PR 19-JUL-1999; 990S-0144434.
PR 20-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144532.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145086.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 05-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 06-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 14-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159429.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.

PR	21-APR-1999;	9905-01304447;
PR	23-APR-1999;	9905-01305119;
PR	23-APR-1999;	9905-01305191;
PR	28-APR-1999;	9905-01314429;
PR	28-APR-1999;	9905-01314499;
PR	30-APR-1999;	9905-01320408;
PR	30-APR-1999;	9905-01324207;
PR	04-MAY-1999;	9905-01324484;
PR	05-MAY-1999;	9905-01324854;
PR	06-MAY-1999;	9905-01324866;
PR	06-MAY-1999;	9905-01324877;
PR	07-MAY-1999;	9905-01328663;
PR	11-MAY-1999;	9905-01328256;
PR	14-MAY-1999;	9905-01342118;
PR	14-MAY-1999;	9905-01342119;
PR	14-MAY-1999;	9905-01342219;
PR	14-MAY-1999;	9905-01342221;
PR	14-MAY-1999;	9905-01343700;
PR	16-MAY-1999;	9905-01343768;
PR	19-MAY-1999;	9905-01349444;
PR	20-MAY-1999;	9905-01351123;
PR	21-MAY-1999;	9905-01353453;
PR	24-MAY-1999;	9905-01356629;
PR	24-MAY-1999;	9905-01360291;
PR	25-MAY-1999;	9905-01363921;
PR	27-MAY-1999;	9905-01367882;
PR	28-MAY-1999;	9905-01372722;
PR	01-JUN-1999;	9905-01375208;
PR	03-JUN-1999;	9905-01375208;
PR	04-JUN-1999;	9905-01377224;
PR	07-JUN-1999;	9905-01380994;
PR	10-JUN-1999;	9905-01385400;
PR	10-JUN-1999;	9905-01388477;
PR	14-JUN-1999;	9905-01391119;
PR	16-JUN-1999;	9905-01394522;
PR	16-JUN-1999;	9905-01394533;
PR	17-JUN-1999;	9905-01394922;
PR	18-JUN-1999;	9905-01394954;
PR	18-JUN-1999;	9905-01394956;
PR	18-JUN-1999;	9905-01394577;
PR	18-JUN-1999;	9905-01394588;
PR	18-JUN-1999;	9905-01394589;
PR	18-JUN-1999;	9905-01394600;
PR	18-JUN-1999;	9905-01394611;
PR	18-JUN-1999;	9905-01394621;
PR	18-JUN-1999;	9905-01394663;
PR	18-JUN-1999;	9905-01397500;
PR	18-JUN-1999;	9905-01397633;
PR	21-JUN-1999;	9905-01398177;
PR	22-JUN-1999;	9905-01398699;
PR	23-JUN-1999;	9905-01403533;
PR	23-JUN-1999;	9905-01403544;
PR	23-JUN-1999;	9905-01403554;
PR	24-JUN-1999;	9905-01406955;
PR	28-JUN-1999;	9905-01408233;
PR	29-JUN-1999;	9905-01409591;
PR	30-JUN-1999;	9905-01428033;
PR	01-JUL-1999;	9905-01412877;
PR	01-JUL-1999;	9905-01418442;
PR	01-JUL-1999;	9905-01421554;
PR	02-JUL-1999;	9905-01420555;
PR	06-JUL-1999;	9905-01423900;
PR	08-JUL-1999;	9905-01428033;
PR	09-JUL-1999;	9905-01429200;
PR	12-JUL-1999;	9905-01429777;
PR	13-JUL-1999;	9905-01435442;
PR	14-JUL-1999;	9905-01436244;
PR	15-JUL-1999;	9905-01440055;
PR	16-JUL-1999;	9905-01440855;
PR	16-JUL-1999;	9905-01440866;
PR	19-JUL-1999;	9905-01443255;
PR	19-JUL-1999;	9905-01443333;
PR	19-JUL-1999;	9905-01443334;
PR	19-JUL-1999;	9905-01443335;

RESULT 33	
ANY70096	
ID	ANY70096 standard; Protein; 152 AA.
XX	
XX	
AC	ANY70096;
XX	
ET	05-JUN-2000 (first entry)
DE	
XX	Zea mays Rad6 protein encoded by clone

RESULT 33	
ANY70096	
ID	AAV70096 standard; Protein; 152 AA.
XX	
XX	AAV70096;
XX	
AC	
XX	
DT	05-JUN-2000 (first entry)
XX	
XX	Zea mays Rad6 protein encoded by clone ZmRAD6-2.
DE	
XX	
KW	Rad6: clone ZmRAD6-2; ubiquitin-conjugating activity; F2; DNA repair;
KW	Ubiquitin Conjugating Enzyme; Ubc; recombination; gene targeting;
KW	male sterility; detection; screen; molecular marker; genotyping; maize.
XX	
OS	Zea mays.
XX	
PN	WC2000C9723-AA..
XX	
PD	24-FEB-2000.
XX	
PF	10-AUG-1999; 99WO-US18126.
XX	
PR	14-AUG-1998; 98US-0096546.
XX	

PA (PION-) PIONEER HI-BRED INT INC.
 XX Mahajan PB;
 XX WPI: 2000-224352/19.
 DR N-PSDB: AAZ51092.
 XX
 PT New nucleic acid encoding Rad6 genes from maize, used to produce
 XX transgenic plants with male sterility and to affect gene targeting
 XX
 PS Claim 1a: Page 77: 85pp: English.
 XX
 CC The present sequence is that of maize Rad6 protein encoded by cDNA clone
 CC ZMRAD6-2. Rad6 is also designated as Ubiquitin Conjugating Enzyme (UBC)
 CC or E2. It has ubiquitin-conjugating activity and is involved in
 CC DNA repair and recombination. Vector containing Rad6 nucleotide sequence
 CC can be used to transform plants to modulate Rad6 expression levels,
 CC affect gene targeting and induce male sterility. The DNA sequence may
 CC be used as probes or primers for detection, quantification or isolation
 CC of gene transcripts; detection of mutations and allelic variants, for
 CC monitoring up-regulation of expression in screening assays and as
 CC molecular markers for genotyping.
 XX
 SO Sequence 152 AA:
 Query Match 12.1% Score 159; DB 27; Length 152;
 Best Local Similarity 29.3%; Pref. No. 4.4e-08;
 Matches 43; Conservative 27; Mismatches 57; Indels 20; Gaps 4;
 QY 9 PEGIMVKTFFDRMDLFSALIKGPIPTPYEDGILYFD:QLPNITPAVPPHFCYLSQSGRI 68
 DB 21 PAGESCAFYDNNIM:WNAVIGGDDIPWGGGIFKLTQCTEDYPNKPIVRFVSR---MF 77
 QY 63 NNULDNGKVCVSLGTWIGKGTERT:SKSLQVLISIQGLII---VNEPYNEAGFDS 125
 DB 78 HPIVYADGSLCDLTL-----QNKSPDIYDVAAL:ISQSLICDNPSPANSRAG--- 127
 QY 126 DRSLQSVSYENRCYNEMALIRVQSMI 152
 DB 128 -----RMESKREYNKRVREWEQSWT 150
 RESULT 34
 AB863736
 ID AB863736 standard; Protein: 150 AA.
 XX
 AC AB863736;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 18000.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US9231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li FWD, Myers EW;
 XX
 DR WPI: 2001-456860/75.
 XX
 DR N-PSDB: ABIC7839.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 18000; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB857737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 190 AA;
 Query Match 12.1% Score 158.5; DB 22; Length 190;
 Best Local Similarity 28.0%; Pref. No. 7e-08;
 Matches 44; Conservative 34; Mismatches 56; Indels 23; Gaps 6;
 QY 4 LATSLEPGIMVKTFFDRMDLFSALIKGPIPTPYEDGILYFD:QLPNITPAVPPHFCYLSQ 63
 DB 7 METTPEGIKVLINESVDIDIGALIDGAGTPYAGIETRVKLIILNKDFFPLTPPKAYFLTK 66
 QY 64 CSGLRNPNLYDNGKVCVSLGTWIGKGTERT:SKSLQVLISIQGLII---VNEPYNEAGF 123
 DB 67 ----PHPNVAANGELCVNTL-----KKDKPDI:GKHITLTKCLLIVPND---ESAL 113
 QY 124 DSDRG--LQEGSYENRCYNEMALIRVQSMI 158
 DB 114 NEERAGKMLIERVDD---YSQRA-----RMWTEHAQP 142
 RESULT 35
 AAR79654
 ID AAR79654 standard; Protein: 152 AA.
 XX
 AC AAR79654;
 XX
 DT 06-DEC-1995 (first entry)
 XX
 DE Human E2 protein.
 XX
 KW Ubiquitin-conjugating enzyme; E2 protein; cell cycle;
 XX cell proliferation; cancer; psoriasis; fibrosis.
 XX
 OS Homo sapiens.
 XX
 XX W09518974-A.
 XX
 PD 13-JUL-1995.
 XX
 PF 04-JAN-1995; 95WO-US00164.
 XX
 PR 13-SEP-1994; 94US-0305520.
 XX
 PR 04-JAN-1994; 94US-0176937.
 XX
 PR 23-MAY-1994; 94US-0247904.
 XX
 PR 27-MAY-1994; 94US-0250795.
 XX
 PA (MITO-) MITOTIX INC.
 XX
 PI Cottarel G, Draetta G, Eckstein JW, Gyris J, Rolfe M;
 XX
 DR WPI: 1995-255137/33.
 XX
 DR N-PSDB: AAQ97844.
 XX
 PT Identifying inhibitors of ubiquitin mediated proteolysis of cell cycle
 PT regulatory proteins - also new ubiquitin conjugating enzymes, their
 PT related nucleic acid, vectors, antibodies etc., useful for regulating

PT e.g. cell proliferation
 PS Disclosure: Page 99-100; 157pp; English.
 XX
 CC Human E2 cDNA (given in AAQ97844) was amplified from a HeLa cell
 CC cDNA library using the primers given in AAQ97842-43. The gene
 CC was subcloned into a baculovirus or pEX vector for expression of
 CC recombinant E2 in Sf9 insect or E. coli cells for use as a
 CC component of an in vitro ubiquitin conjugating system.
 XX
 SQ Sequence 152 AA;
 Query Match 11.9%; Score 157; DB 16; Length 152;
 Best Local Similarity 30.3%; Pred. No. 7.1e-08;
 Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;
 QY 9 PEGIMVKTEDMDLFSALIKGPTR:PYHEDGLYFDIQPNLYAPVPHFCYLSQSGRL 68
 DB 21 PVSVSGAPSENNIMQWNAVIRGEGTPEEDGTFKLVIEFSEYPNKPTVRLSK---MF 77
 QY 69 NPRLYDNGKVCVSLIGTWIGKTERWTSKSLIQVLISIQGLILVNPYPYNEAGFDSRG 128
 DB 78 HPNVYADGSCLDIL-----QNRWSPYDVSSILTSIQSL--DEPNNSPA---NSQA 126
 QY 129 LQEGYENSRCYNEALIRVQS 150
 DB 127 AQIYQENKREYKRYSAIVQS 148
 RESULT 36
 AAQ39567
 ID AAY19967 standard; Protein: 152 AA.
 AC AAY3967;
 DT 15-DEC-1999 (first entry)
 DE Human rad6 homolog protein sequence.
 KW ubiquitin conjugating enzyme; UbC6; ubiquitin-mediated proteolysis;
 KW cell cycle regulatory protein; ubiquitination inhibitor; atherosclerosis;
 KW proliferative disorder; cancer; restenosis; tissue connective disorder;
 KW wound healing; fibrosis disorder; rheumatoid arthritis; scleroderma;
 KW insulin dependent diabetes mellitus; glomerulonephritis; cirrhosis;
 KW diagnosis; therapy; rad6; ss.
 QS Homo sapiens.
 XX US5968762-A.
 XX 19-OCT-1999.
 XX 07-JUN-1995; 95US-0486663.
 XX 04-JAN-1994; 94US-0176957.
 XX 23-MAY-1994; 94US-0247904.
 XX 27-MAY-1994; 94US-0250795.
 XX 13-SEP-1994; 94US-0305520.
 XX (MITO-) MITOTIX INC.
 XX Chiu MI, Cottarel G, Berlin V, Damaguez V, Draetta G, Rolfe M;
 WPJ: 1999 590402/50.
 DR N-PSDB: AA27567.
 XX Identifying ubiquitination inhibitors using novel ubiquitin conjugating
 PT enzymes -
 XX Example 2: Column 89-90; 61pp; English.
 PS This sequence is the human rad6 protein. The invention relates to
 CC assays for identifying an inhibitor of ubiquitin-mediated proteolysis of

CC a cell-cycle regulatory protein comprising contacting a candidate agent
 CC with an ubiquitin-conjugating system and measuring the level of
 CC ubiquitination. The ubiquitin-conjugating system comprises:
 CC (a) a reconstituted protein mixture including a ubiquitin conjugating
 CC enzyme (UbC6) produced by the expression of a nucleic acid which
 CC hybridizes under high stringency conditions to human UbC6, Candida
 CC albicans UbC6, or Schizosaccharomyces pombe UbC6 coding sequences;
 CC (b) a regulatory protein; and (c) ubiquitin. The polynucleotides are
 CC useful for identifying ubiquitination inhibitors. The polynucleotides,
 CC polypeptides, antisense compounds and antibodies against them may also be
 CC useful for the treatment and/or diagnosis of proliferative disorders
 CC (e.g. cancer, atherosclerosis, or restenosis), tissue connective
 CC disorders, controlling wound healing, and disorders characterized by
 CC fibrosis (e.g. rheumatoid arthritis, insulin dependent diabetes mellitus,
 CC glomerulonephritis, cirrhosis, and scleroderma).
 XX
 SQ Sequence 152 AA;
 Query Match 11.9%; Score 157; DB 20; Length 152;
 Best Local Similarity 30.3%; Pred. No. 7.1e-08;
 Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;
 QY 9 PEGIMVKTEDMDLFSALIKGPTR:PYHEDGLYFDIQPNLYAPVPHFCYLSQSGRL 68
 DB 21 PVSVSGAPSENNIMQWNAVIRGEGTPEEDGTFKLVIEFSEYPNKPTVRLSK---MF 77
 QY 69 NPRLYDNGKVCVSLIGTWIGKTERWTSKSLIQVLISIQGLILVNPYPYNEAGFDSRG 128
 DB 78 HPNVYADGSCLDIL-----QNRWSPYDVSSILTSIQSL--DEPNNSPA---NSQA 126
 QY 129 LQEGYENSRCYNEALIRVQS 150
 DB 127 AQIYQENKREYKRYSAIVQS 148
 RESULT 37
 AAB03175
 ID AAB03175 standard; Protein: 152 AA.
 AC AAB03175;
 DT 23-OCT-2000 (first entry)
 DE Human rad6 homologue UbC2.
 KW Human UbC2; rad6 homologue; ubiquitin conjugating enzyme;
 KW ubiquitin mediated proteolysis; human; cellular protein half life;
 KW ubiquitination inhibitor; p53; cyclin; cell cycle regulator;
 KW myc deregulation; human papillomavirus; HPV-18 E6 protein;
 KW cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;
 KW psoriasis; connective tissue disorder; wound healing; cytostatic;
 KW antiproliferative; anticancer; antipsoriatic.
 XX Homo sapiens.
 XX US6068982-A.
 XX 30-MAY-2000.
 XX 17-DEC-1996; 96US-0767942.
 XX 07-JUN-1995; 95US-0486663.
 XX 04-JAN-1994; 94US-0176937.
 XX 23-MAY-1994; 94US-0247904.
 XX 27-MAY-1994; 94US-0250795.
 XX 13-SEP-1994; 94US-0305520.
 XX (MITO-) MITOTIX INC.
 XX Chiu MI, Cottarel G, Berlin V, Draetta G, Damaguez V, Rolfe M;
 WPJ: 2000-410654/35.
 DR N-PSDB: AAB61622.

XX identifying an inhibitor of ubiquitin mediated proteolysis of
 PT regulatory protein for treating cancers involves measuring
 PT ubiquitination levels of the protein in the presence of candidate agent
 PT in an eukaryotic cell
 XX

PS Example 2: Column 95-96: 73pp; English.

XX The invention relates to a method of identifying an inhibitor of
 CC ubiquitin mediated proteolysis of a cell cycle regulatory protein
 CC comprising contacting an engineered eukaryotic cell with a candidate
 CC agent. The eukaryotic cells is engineered to express a recombinant
 CC human, Candida albicans or Schizosaccharomyces pombe ubiquitin-
 CC conjugating enzyme (AA03169-B03171), a cell cycle regulatory protein
 CC (such as p53) and ubiquitin. The specification also discloses novel
 CC Candida albicans and Schizosaccharomyces pombe ubiquitin-conjugating
 CC enzymes, caUCE and spUCE (AA03170, AA03171), and two novel human
 CC ubiquitin-conjugating enzymes, huBCE and rapBCE (AA03169, AA03173).
 CC The ubiquitin-mediated proteolysis system is the major pathway for the
 CC selective, controlled degradation of intracellular proteins in
 CC eukaryotic cells, and is important in controlling the half-lives of
 CC cellular proteins, and in particular, this system controls the levels of proteins
 CC involved in cell cycle progression. Alterations in the ubiquitination of
 CC these proteins may therefore play a role in the development of cancers.
 CC For example, human papillomaviruses such as HPV-18 encode a transforming
 CC protein, E6 (AA03176), which combines with a cellular E6-associated
 CC protein (E6-AP; AA03177) to stimulate the ubiquitination of p53, thus
 CC targeting it for degradation. The ubiquitination inhibitors identified
 CC according to the method of the invention are useful for treatment of
 CC cervical cancers and connective tissue disorders and for controlling the
 CC wound healing process. They are also useful in treatment of hyperplastic
 CC epidermal conditions such as psoriasis, neoplastic epidermal conditions,
 CC skin cancers e.g., basal cell carcinomas, squamous cell carcinomas. The
 CC inhibitors are useful for deregulating myc expression and rendering the
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-
 CC mediated degradation of cyclins are useful as antiproliferative agents.
 CC The present sequence represents the human rad6 homologue UBC2.

XX Sequence 152 AA:

Query Match 11.9%; Score 157; DB 21; Length 152;
 Best Local Similarity 30.3%; Pred. No. 7.1e-08;
 Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;

QY 9 PEGIMVKITFEDRMDFSAALIKGTRTPYEDGLYLFDFQLENIYPAVPPHFCYLSQSGRL 68

DB 23 PVGVSCAPSENNIMQWNAVIFGPEGTPEDCFKLVIEFSEYFNKPPVPLSK---MF 77

QY 69 NPNLYDNKVCVVSLLGTWICKGTERTWSKSLQLVLISQGLILVNEPYNAGFDSORG 128

DE 78 HFNVAIXSICIDIL-----QNRWSPYDVSSILTSIQSL--DEPNPNSPA--NSQA 126

QY 129 LQESVENSRCYNMAIRVVQS 150

DB 127 AQHYNKKRYEKRVSAIVQS 149

RESULT 38

ABP42383

10 ABP42383 standard; Protein: 166 AA.

AC ABP42383;

XX 22-AUG-2002 (first entry)

DE Human ovarian antigen HNORH05, SEQ ID NO:3515.

XX Human: ovarian antigen: ovary; ovarian: breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumor; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;

KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.

XX Homo sapiens.

OS WO200200677-A1.

PN 03-JAN-2002.

XX 07-JUN-2001: 2001WO-US18569.

XX 07-JUN-2000: 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Hirse CH, Rosen CA;

XX WPI: 2002-147878/19.

DR N-PSDB; ABQ55460.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases.

XX Claim 11: SEQ ID No 3515; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (AA041054-
 CC ABP43228) and to cDNAs encoding them (AB054131-AA056305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pcl_sequences.

XX Sequence 166 AA:

Query Match 11.9%; Score 157; DB 23; Length 166;

Best Local Similarity 30.3%; Pred. No. 8.1e-08;

Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;

QY 9 PEGIMVKITFEDRMDFSAALIKGTRTPYEDGLYLFDFQLENIYPAVPPHFCYLSQSGRL 68

DB 23 PVGVSCAPSENNIMQWNAVIFGPEGTPEDCFKLVIEFSEYFNKPPVPLSK---MF 91

QY 69 NPNLYDNKVCVVSLLGTWICKGTERTWSKSLQLVLISQGLILVNEPYNAGFDSORG 128

DB 72 HFNVAIXSICIDIL-----QNRWSPYDVSSILTSIQSL--DEPNPNSPA--NSQA 140

QY 129 LQGYENSRYNEMALIRVQS 150
 DB 141 AQLYQENKREYKRVSAIVEQS 162

RESULT 39
 AAB43423

ID AAB43423 standard; Protein: 196 AA.
 AC AAB43423:
 XX 08-FEB-2001 (first entry)
 XX human cancer associated protein sequence SEQ ID NO:868.
 XX Human: Cancer associated gene; cancer antigen; detection; cancer;
 KW diagnostic; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antisthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 CS Homo sapiens.
 XX
 FN W0203055350-A1.
 XX
 PD 21-SEP-2002.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 FA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen SM, Ruben SM;
 XX
 DR WPI: 2000-587553/55.
 DR N-PSDB: AAC77632.
 XX
 PI Note: Isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer.
 XX
 PS Claim 11; Page 1423; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43396 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antisthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiac; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 196 AA;

Query Match 11.9%; Score 157; DB 21; Length 196;
 Best Local Similarity 30.3%; Pred. No. 1e-07;
 Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;

QY 9 PEGIMVKITFEDRMDFLSALIKGPTKIPYEDCLYLFDLQLPNTYPAVPPHFCYISQCSGRI 68
 DB 65 PVGVSGAPSENNIMQNAVIFGPGTFEDGTEKLVIEFSEYYPKRPPTVFLSK---MF 121
 QY 69 NPNIYDNGKVCVSLQGTWTKGTERWT SKSLQVLIISIGLILVNEPYNFAFGFSDRKG 128
 DB 122 HPVYVAGSICLDIL-----QNRWSPTYGVSSILTSQSIL--DEPNPNSPA--NSQA 170
 QY 129 LQGYENSRYNEMALIRVQS 150
 DB 171 AQLYQENKREYKRVSAIVEQS 192

RESULT 40
 AAG73872
 ID AAG73872 standard; Protein: 198 AA.
 AC AAG73872:
 XX 03-SEP-2001 (first entry)
 XX Human colon cancer antigen; protein SEQ ID NO:4636.
 DE Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 5.
 XX
 CS Homo sapiens.
 XX
 FN W0200122920-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI: 2001-235357/24.
 DR N-PSDB: AAB33303.
 XX
 PI Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PI useful for preventing, diagnosing and/or treating colorectal cancers.
 XX
 PS Claim 11; Page 6439-6440; 9803pp; English.
 XX
 CC AAB32943 to AAB37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate p
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps.
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAB37196 to AAB37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX

```

SQ Sequence 198 AA;
Query Match      11.9%; Score 157; DB 22; Length 198;
Best local similarity 30.3%; Pred. No. 11e-07;
Matches 43; Conservative 25; Mismatches 60; Indels 14; Caps 4;

Qy 9 PEGIMVKTEDRMFLSALIKGTETPYEDGLYLFDTQJPNIIYPVPPHFCYLSQCSGL 68
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 67 PVGVSCAPSENNINQWNAVIEGEGIPFEGGIFKIVIERSEEPNKPPTVRFLSK---MF 123
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 69 NENLYDNCKKVCVSLGHWICKGTERWTSSLLQVLISLQGLILVNEPYNNAGFSDRG 128
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 124 HPRVYADGSLCLDIL-----GNRMSPTDYVSSILISIOSLL--DEPNPSPA--NSQA 172
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 129 LQGVNSKCYNEMALLKVVCS 150
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 173 AGLYQENKREYKRVSAIVEQS 194
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

Search completed: April 10, 2003, 10:32:12
Job time : 52.2353 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:27:54 ; Search time 12.3529 Seconds
(without alignments)
1945.578 Million cell updates/sec

Title: US-99-930-026-1

Perfect score: 1315

Sequence: I M A I A I S L P E G I M V K T F E D R P G E A S Q G S D S E C C A Q G L A P S 250

Scoring table: BIOSUM62

Gapop 10.0 , Gapext: 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum

Post-processing: Minimum March 08
Maximum March 1008

Maximum Match 100%
Listing first 135 summaries

1544

```
Database :
          PIR_13:
          1: pir:*
```

* : Caid
 * : Irid
 : C
 : I

2: 2
3: 3
4: 4
5: 5
6: 6
7: 7
8: 8
9: 9
10: 10
11: 11
12: 12
13: 13
14: 14
15: 15
16: 16
17: 17
18: 18
19: 19
20: 20
21: 21
22: 22
23: 23
24: 24
25: 25
26: 26
27: 27
28: 28
29: 29
30: 30
31: 31
32: 32
33: 33
34: 34
35: 35
36: 36
37: 37
38: 38
39: 39
40: 40
41: 41
42: 42
43: 43
44: 44
45: 45
46: 46
47: 47
48: 48
49: 49
50: 50
51: 51
52: 52
53: 53
54: 54
55: 55
56: 56
57: 57
58: 58
59: 59
60: 60
61: 61
62: 62
63: 63
64: 64
65: 65
66: 66
67: 67
68: 68
69: 69
70: 70
71: 71
72: 72
73: 73
74: 74
75: 75
76: 76
77: 77
78: 78
79: 79
80: 80
81: 81
82: 82
83: 83
84: 84
85: 85
86: 86
87: 87
88: 88
89: 89
90: 90
91: 91
92: 92
93: 93
94: 94
95: 95
96: 96
97: 97
98: 98
99: 99
100: 100
101: 101
102: 102
103: 103
104: 104
105: 105
106: 106
107: 107
108: 108
109: 109
110: 110
111: 111
112: 112
113: 113
114: 114
115: 115
116: 116
117: 117
118: 118
119: 119
120: 120
121: 121
122: 122
123: 123
124: 124
125: 125
126: 126
127: 127
128: 128
129: 129
130: 130
131: 131
132: 132
133: 133
134: 134
135: 135
136: 136
137: 137
138: 138
139: 139
140: 140
141: 141
142: 142
143: 143
144: 144
145: 145
146: 146
147: 147
148: 148
149: 149
150: 150
151: 151
152: 152
153: 153
154: 154
155: 155
156: 156
157: 157
158: 158
159: 159
160: 160
161: 161
162: 162
163: 163
164: 164
165: 165
166: 166
167: 167
168: 168
169: 169
170: 170
171: 171
172: 172
173: 173
174: 174
175: 175
176: 176
177: 177
178: 178
179: 179
180: 180
181: 181
182: 182
183: 183
184: 184
185: 185
186: 186
187: 187
188: 188
189: 189
190: 190
191: 191
192: 192
193: 193
194: 194
195: 195
196: 196
197: 197
198: 198
199: 199
200: 200
201: 201
202: 202
203: 203
204: 204
205: 205
206: 206
207: 207
208: 208
209: 209
210: 210
211: 211
212: 212
213: 213
214: 214
215: 215
216: 216
217: 217
218: 218
219: 219
220: 220
221: 221
222: 222
223: 223
224: 224
225: 225
226: 226
227: 227
228: 228
229: 229
230: 230
231: 231
232: 232
233: 233
234: 234
235: 235
236: 236
237: 237
238: 238
239: 239
240: 240
241: 241
242: 242
243: 243
244: 244
245: 245
246: 246
247: 247
248: 248
249: 249
250: 250
251: 251
252: 252
253: 253
254: 254
255: 255
256: 256
257: 257
258: 258
259: 259
260: 260
261: 261
262: 262
263: 263
264: 264
265: 265
266: 266
267: 267
268: 268
269: 269
270: 270
271: 271
272: 272
273: 273
274: 274
275: 275
276: 276
277: 277
278: 278
279: 279
280: 280
281: 281
282: 282
283: 283
284: 284
285: 285
286: 286
287: 287
288: 288
289: 289
290: 290
291: 291
292: 292
293: 293
294: 294
295: 295
296: 296
297: 297
298: 298
299: 299
300: 300
301: 301
302: 302
303: 303
304: 304
305: 305
306: 306
307: 307
308: 308
309: 309
310: 310
311: 311
312: 312
313: 313
314: 314
315: 315
316: 316
317: 317
318: 318
319: 319
320: 320
321: 321
322: 322
323: 323
324: 324
325: 325
326: 326
327: 327
328: 328
329: 329
330: 330
331: 331
332: 332
333: 333
334: 334
335: 335
336: 336
337: 337
338: 338
339: 339
340: 340
341: 341
342: 342
343: 343
344: 344
345: 345
346: 346
347: 347
348: 348
349: 349
350: 350
351: 351
352: 352
353: 353
354: 354
355: 355
356: 356
357: 357
358: 358
359: 359
360: 360
361: 361
362: 362
363: 363
364: 364
365: 365
366: 366
367: 367
368: 368
369: 369
370: 370
371: 371
372: 372
373: 373
374: 374
375: 375
376: 376
377: 377
378: 378
379: 379
380: 380
381: 381
382: 382
383: 383
384: 384
385: 385
386: 386
387: 387
388: 388
389: 389
390: 390
391: 391
392: 392
393: 393
394: 394
395: 395
396: 396
397: 397
398: 398
399: 399
400: 400
401: 401
402: 402
403: 403
404: 404
405: 405
406: 406
407: 407
408: 408
409: 409
410: 410
411: 411
412: 412
413: 413
414: 414
415: 415
416: 416
417: 417
418: 418
419: 419
420: 420
421: 421
422: 422
423: 423
424: 424
425: 425
426: 426
427: 427
428: 428
429: 429
430: 430
431: 431
432: 4

3: dir3: \star
4: dir4: \star

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			ID	Description
	Score	% Match	Length		
1	474	36.2	1132	2	H84545
2	424.5	32.2	543	2	G90570
3	392	29.7	182	2	S84749
4	372	28.3	4845	2	I31067
5	269	20.5	423	2	I15350
6	179	13.6	172	2	A21906
7	169	12.9	247	2	I42856
8	165.5	12.6	165	2	S28951
9	164	12.5	153	2	C86304
10	161	12.2	153	2	S58992
11	159	12.1	151	2	S13529
12	159	12.1	205	2	I32559
13	156.5	12.1	151	2	T51931
14	157	11.9	151	2	T45220
15	157	11.9	152	2	I51913
16	157	11.9	152	2	A42436
17	157	11.9	152	2	B41222
18	156.5	11.9	166	2	T43235
19	153	11.6	152	2	A41222
20	152.5	11.6	151	2	S73430
21	151	11.5	151	2	A39392
22	151	11.5	163	2	H96818
23	151	11.5	233	1	S19158
24	150	11.4	154	2	E86484
25	148	11.3	298	2	A49630
26	145.5	11.1	199	2	T33629
27	144.5	11.0	152	2	S43781
28	144	11.0	148	2	T17532
29	143.5	10.9	152	2	S43783

Query Match	12.5%	Score 164;	DB 2;	Length 153;
Best Local Similarity	27.6%	Pred. No. 3.5e-07;		
Matches 42:	Conservative 28;	Mismatches 62;	Indels 20;	Gaps

Qy 123 FUSDRLQEGYNSRCYNEMALRVVQSMTCQ 154
 . | ; || : | : | : |

Dd 79 PHPNYKDGTVICISILHAGDGDPPNNFYESSRSPVQSVKEIKILLSVMSLA--EP-NDPS 135

Qy 122 GFSDS 126
 .: :
Db 136 GANID 140

RESULT 19
A41222
ubiquitin-protein ligase (EC 6.3.2.19) E2A - human
N:Alternate names: ubiquitin-conjugating enzyme HHR6A
C:Species: Homo sapiens (man)
C>Date: 19-Jun-1992 #sequence_revision 18-Jun-1992 #text_change 03-Jun-2002
C:Accession: A41222
R:Koken, M.H.M.; Reynolds, P.; Jaspers-Dekker, I.; Prakash, L.; Prakash, S.; Bootsma
Proc. Natl. Acad. Sci. U.S.A. 88, 8865-8869, 1991
A>Title: Structural and functional conservation of two human homologs of the yeast D
A:Reference number: A41222; PMID:92020951; PMID:1717990
A:Accession: A41222
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-152 <KK>
A:Cross-references: GB:M74524; NID:g184043; PIDN:AAA35981.1; PID:g184044
C:Genetics:
A:Gene: GDB-UHK2A; UBC2; HHRGA
A:Cross-references: GDB:l31647; OMIM:312180
A:Map position: Xq24-Xq25
C:Superfamily: human ubiquitin-protein ligase E2
C:Keywords: Ligase; protein degradation; ubiquitination
F:88/Active site: Cys #status predicted

Query Match 11.6%; Score 153; DB 2; Length 152;
Best Local Similarity 28.9%; Pred.No. 3.6e-06;
Matches 41; Conservative 27; Mismatches 60; Indels 14; Gaps 4;

Qy 9 PEGIMVKTFEDRMDELASLTKGTPRTPYEDGLYLFDIQLPNIYPAPVPHPFCVLSCSGRL 68
 | | | | | . : : : : | | | | | : : : : : | | | | | : : : : :
Db 21 PAGVSgapSENMMVNNAVIFCGPTGCDCGTFKLTIERTEEYPKNKPVTVEFSK---MF 77

Qy 69 NPNIYNCKVCVVLLTATGKTGERMTSKSSLSIOGLINPEYNEAFGDSDRG 128
 | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 78 HPNYVDGSIGCIDLI-----QNWSPTYDVSYSLTSIQSI-L--DEPNNSPA--NSQA 126

Qy 129 LOEGYSNRCYNEMALIRVVQS 150
 | | | | | . : : : : | | | | | : : : : : | | | | | : : : : :
Db 127 AQIVENKRNEYEKVRSAIVEQS 148

RESULT 20
S71430
DNA repair protein mus-8 - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jul-2000
C:Accession: S71430
R:Soshi, T.; Sakuraba, Y.; Kaefer, F.; Inoue, H.
Curr. Genet. 30, 224-231, 1996
A>Title: The mus-8 gene of Neurospora crassa encodes a structural and functional hom
A:Reference number: S71430; PMID:96337904; PMID:8753651
A:Accession: S71430
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <SOS>
A:Cross-references: EMBL:D78472; NID:g150728; PIDN:BAALL380.1; PID:g1100729
A>Note: the authors translated the codon GCC for residue 26 as Ser
C:Genetics:
A:Gene: mus-8
A:Introns: 14/3; 36/2; 111/2
C:Superfamily: human ubiquitin-protein ligase R2

Query Match 11.6%; Score 152.5; DB 2; Length 151;
Best Local Similarity 27.5%; Pred.No. 3.6e-06;
Matches 38; Conservative 24; Mismatches 55; Indels 21; Gaps 4;

QY 4 LATSLEPGIMVKTERRMDFSAIKIGKTRTPYEDGLYFDIQLPNIYVAVPPHFCYLSQ 63
 DB 16 MOTDPPAIVASASVPDQNVNMTAVNVAIIGPADTFFEDG*FRLVMHFQYVKNPKPSVKFEISE 75
 QY 64 QSGRLNPNLYDNGKVCVSLGTWIGKTERMTSKSSLLQVLIS:QGLLILVNPYYNEAGF 123
 DB 76 ---MHPRVATSGELGLDIL-----QNRWSPDYVAVALTISQSL--NDP----- 116
 QY 124 NSDRGLQVGYNSRCYNE 141
 DB 127 --NIGSPANVAVASNYKO 132
 RESULT 21
 A3392
 RA16 DNA repair homolog chr6 - fruit fly (*Drosophila melanogaster*)
 C:Species: *Drosophila melanogaster*
 C>Date: 26-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 20-Sep-1999
 C:Accession: A3392
 R:Koken, M.; Reynolds, P.; Bootsma, D.; Hoeijmakers, J.; Prakash, S.; Prakash, L.
 Proc. Natl. Acad. Sci. U.S.A. 88, 3832-3836, 1991
 A:Title: chr6, a *Drosophila* homolog of the yeast DNA-repair gene RAD6.
 A:Reference number: A3392; MUID:91219466; PMID:1302572
 A:Accession: A3392
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-151 <KOK>
 A:Cross-references: GB:M63791; OH:M63792; NID:g156731; PID:g156732
 C:Genetics:
 A:Gene: Flybase:fb06
 A:Cross-references: Flybase:Flybase004436
 C:Superfamily: human ubiquitin-protein ligase E2
 Query Match 11.5%; Score 151; DB 2; Length 151;
 Best Local Similarity 28.2%; Pred. No. 4.9e-06;
 Matches 40; Conservative 30; Mismatches 58; Indels 14; Gaps 4;
 QY 9 PEGIMVKTERRMDFSAIKIGKTRTPYEDGLYFDIQLPNIYVAVPPHFCYLSQCSGRL 68
 DB 21 PTVGSAVPTDNTNMIWNAVIGPHDIPEDGTFKLTETETPKPTVREKVV---F 77
 QY 69 NPNLYDNGKVCVSLGTWIGKTERMTSKSSLLQVLISQGLLILVNPYYNEAGFDSRG 128
 DB 78 HENYADSGICLDL-----QNRWSPDYVAVALTISQSL--SDPNPASPANSTAAQ 128
 QY 123 LQGVYNSRCYNEAMILRYVQS 150
 DB 129 LYK--ENRREYKRVACVEQS 148
 RESULT 22
 B96818
 hypothetical protein F9K20.8 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: B96818
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
 Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B96818
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-154 <STO>
 A:Cross-references: GB:AE005173; NID:g3834310; PIDN:AA83026.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: human ubiquitin-protein ligase E2

A:Gene: F9K20.8
 A:Map position: 1
 C:Superfamily: human ubiquitin-protein ligase E2
 Query Match 11.5%; Score 151; DB 2; Length 163;
 Best Local Similarity 31.3%; Pred. No. 5.4e-06;
 Matches 36; Conservative 24; Mismatches 43; Indels 12; Gaps 4;
 QY 3 LLATSLEPGIMVKTERRMDFSAIKIGKTRTPYEDGLYFDIQLPNIYVAVPPHFCYLS 62
 DB 17 LLESEAP-GTASPSSENNMYNNMILGPTGSPYGGVFKLELPEEYPMAPKVRLELT 75
 QY 63 QCSGRNPNLYDNGKVCVSLGTWIGKTERMTSKSSLLQVLISQGLLILVNEP 116
 DB 76 KZ---YHPNIDKIGRICLDL-----KQKWSPALQIRTVLLRSTQALLSAPNP 120
 RESULT 23
 S19158
 ubiquitin-protein ligase (EC 6.3.2.19) E2 - African swine fever virus (isolate Malawi)
 C:Species: African swine fever virus, ASFV
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 03-Jun-2002
 C:Accession: S19158
 R:Hindamp, P.M.; Arnold, J.E.; Mayer, R.J.; Dixon, L.K.
 EMBO J. 11, 361-366, 1992
 A:Title: A ubiquitin conjugating enzyme encoded by African swine fever virus.
 A:Reference number: S19158; MUID:92155177; PMID:1310934
 A:Accession: S19158
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-213 <HIN>
 A:Cross-references: EMBL:X62440; NID:g58649; PIDN:CAA44305.1; PID:g58650
 C:Superfamily: African swine fever virus ubiquitin-protein ligase E2
 C:Keywords: ligase; protein degradation
 Query Match 11.5%; Score 151; DB 1; Length 213;
 Best Local Similarity 32.4%; Pred. No. 7.9e-06;
 Matches 36; Conservative 20; Mismatches 43; Indels 12; Gaps 4;
 QY 10 PGIWVKTERRMDFSAIKIGKTRTPYEDGLYFDIQLPNIYVAVPPHFCYLSQCSGRLN 69
 DB 19 EHFISVNEKLTEDWILKGGPDPTLYEGGLFKAKIVFPKPYEPRLTETSE---MWII 75
 QY 70 PNLVDNGKVCVSLGTWIGKTERMTSKSSLLQVLISQGLLILVNEP 116
 DB 76 PNIYSGKLCISILH---GNAPEQGMTHSPAQKIDTVILSV--ISLILNEP 121
 RESULT 24
 E86484
 hypothetical protein F7P23.6 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: E86484
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
 Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86484
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-154 <STO>
 A:Cross-references: GB:AE005172; NID:g6978907; PIDN:AAF4303.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.4%; Score 150; DB 2; Length 154;
 Best Local Similarity 33.3%; Pred. No. 6.1e-06;
 Matches 31; Conservative 23; Mismatches 27; Indels 12; Gaps 3;

QY 24 FSALIKQPRIPYEDGLYFDIQLPNLYPAVPHFCYLSQCSGRINP 83
 DB 39 WTAVIRGPDCTPEGCMFNLSIKFTYDPPKPKFTFKIPTI---YHPINDEGSCMNL 95

QY 84 GTWIGKTERWTSKSLGLVLSIQLIIVNEP 116
 DB 96 -----KDKWTPALWVKKVLLSI--LIIIEKP 119

RESULT 25
 A49630
 ubiquitin conjugating enzyme - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 05-Nov-1999
 C:Accession: A49630
 R:Plon, S.E.; Leppig, K.A.; Do, H.N.; Groudine, M.
 Proc. Natl. Acad. Sci. U.S.A. 90, 10484-10488, 1993
 A:Title: Cloning of the human homolog of the CDC34 cell cycle gene by complementation in
 A:Reference number: A49630; MUID:94068425; PMID:8248134
 A:Accession: A49630
 A:Status: preliminary; translated from CB/PMHL/DBBJ
 A:Molecule type: mRNA
 A:Residues: 1-298 <RES>
 A:Cross-references: GB:222005; NID:q388308; PIDN:AAC97534.1; PID:q388309
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.3%; Score 148; DB 2; Length 298;
 Best Local Similarity 30.3%; Pred. No. 2.9e-05;
 Matches 44; Conservative 28; Mismatches 55; Indels 18; Gaps 7;

QY 20 ECWVKTIFERMDLFS--ALIKSPIRIPYEDGLYFDIQLPNLYPAVPHFCYLSQCSGR 67
 DB 88 EGFVY-TEVDEGDIYHWEVAIFGPNYYKGYFKAKDFPIDYDYSPPAEFRTK---M 143

QY 68 INPXDYNGKVCVSLGLTWICK-----GTRWTSKSLQLVLSIQLIIVNEP-YYNE 120
 DB 144 WEPN:YHICDVQCLSHHPVDVQSGELPSFRWKNPTQNVRTILISV--ISLLNRPNTFSP 201

QY 121 AGFISD---RGLQGYENSRCYNEM 142
 DB 202 ANVDASVWYKKKESKGGREYTDI 226

RESULT 26
 T33629
 hypothetical protein F4069.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T33629
 R:Graves, T.; Suterer, C.; Ozersky, P.
 submitted to the EMBL Data Library, October 1998
 A:Description: The sequence of C. elegans cosmid F4069.
 A:Reference number: 221378
 A:Accession: T33629
 A:Status: preliminary; translated from CB/PMHL/DBBJ
 A:Molecule type: tNA
 A:Residues: 1-199 <GRA>
 A:Cross-references: EMBL:AF099193; PIDN:ACG68796.1; GSPDB:GN00021; CESP:F4069.3
 A:Experimental source: strain Bristol N2; Clone F4069
 C:Genetics:
 A:Gene: CESP:F4069.3
 A:Map position: 3
 A:Introns: 40/3; 72/3; 100/2; 342/3
 C:Superfamily: yeast ubiquitin-protein ligase UBCL

Query Match 11.1%; Score 145.5; DB 2; Length 199;
 Best Local Similarity 29.9%; Pred. No. 2.2e-05;
 Matches 32; Conservative 25; Mismatches 39; Indels 11; Gaps 3;

QY 11 GTMWKTFDRMDLFSALIKSPIRIPYEDGLYFDIQLPNLYPAVPHFCYLSQCSGRINP 70
 DB 26 GIMIEILNENLEIKHIGRPDTPYAGGMFDLKIIPDQYFSPDNWAFSTKI---WHP 82

QY 71 NL-YDNCVKVCVSLGLTWICKGTERWTSKSLQLVLSIQLIIVNEP 116
 DB 83 NVSSGTGVICLDLI-----KDMAASITLITVLISQALMCTPEP 122

RESULT 27
 S4378:
 ubiquitin-conjugating enzyme UBCL - Arabidopsis thaliana
 N:Alternate names: ubiquitin-conjugating enzyme E2
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
 C:Accession: S43781
 R:Sullivan, M.L.; Carpenter, T.B.; Vierstra, R.D.
 Plant Mol. Biol. 24, 651-661, 1994
 A:Title: Homologs of wheat ubiquitin-conjugating enzymes - TaUBC1 and TaUBC4 are en
 A:Reference number: S43781; MUID:94207190; PMID:8155884
 A:Accession: S43781
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-152 <SOL>
 A:Cross-references: EMBL:119352; NID:q431259; PIDN:AAA32897.1; PID:q431260
 C:Genetics:
 A:Gene: UBC1
 A:Introns: 42/2; 51/1; 81/1; 110/3
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.0%; Score 144.5; DB 2; Length 152;
 Best Local Similarity 26.4%; Pred. No. 1.8e-05;
 Matches 42; Conservative 29; Mismatches 53; Indels 35; Gaps 5;

QY 9 PEGIMVKTIFERMDLFSALIKSPIRIPYEDGLYFDIQLPNLYPAVPHFCYLSQCSGR 68
 DB 21 PAGTSGAPQDNNITMNAVIFGPDITPDGDTFKLSIQFSEDPNKPPTVREVSF---MF 77

QY 69 NPNLYDNGKVCVSLGLTWICKGTERWTSKSLQLVLSIQLIIVNEP-YYNEPYNEAGFDS 125
 DB 78 HPIYVAGSGICLDLI-----QNWSPYDYVAALISIQSLICDPNPNPANSEA----- 126

QY 126 DRGLQGYENSRCYNEMALIRVQSMQLVLRPPPEVFFQ 164
 DB 127 -----ARMYSE-----SKREYNRRVRDVEQ 147

RESULT 28
 T37532
 ubiquitin-conjugating enzyme - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T37532
 R:Murphy, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: 221721
 A:Accession: T37532
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-148 <MUR>
 A:Cross-references: EMBL:298595; PIDN:CAB11183.1; GSPDB:GN00066; SPDB:SPAC11E3.04C
 A:Experimental source: strain 972h; cosmid c11E3
 C:Genetics:
 A:Gene: SPDB:SPAC11E3.04C
 A:Map position: 1
 A:Introns: 9/3; 63/3; 92/1
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 11.0%; Score 144; DB 2; Length 148;
 Best Local Similarity 29.6%; Pred. No. 2e-05;
 Matches 32; Conservative 25; Mismatches 41; Indels 10; Gaps 2;

A:Accession: J04894
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-152 <YAM>
A:Cross-references: DDBJ:D81004; NID:gli181557; PIDN:BAAL1675.1; PID:gli181558
C:Genetics:
A:Gene: GDH-URK2N
A:Cross-references: GDB:6953724
C:Superfamily: human ubiquitin-protein ligase E2
C:Keywords: ligase; protein degradation; ubiquitination
E:87/Active site: Cys #status predicted

Query Match 10.4%; Score 137; DB 2; Length 152;
Best Local Similarity 24.5%; Pred. No. 8.4e-05;
Matches 35; Conservative 31; Mismatches 63; Indels 14; Gaps 4;

QY 3 ILATSLPESIMVKTFEDRMDFSAALIKGPTRTPEYDGLYLFDIQLNPYVAPVPHFCYLS 62
DB 15 LIAEVPV-GIARFESNANYEHVIAHQSPFEGGFKLEIFLPEEYMAAPKYRMT 73
QY 63 QCSGRLENLYDNGKVCVSLGWIGKGTERTWTSKSSLLQVLISIQGLIIV---NEPYN 119
DB 74 KT---YHNVNDKLRIGLDL-----KQKSPALQIRTVLLSIQALLSAPNDPLAN 123
QY 120 EACFSDSNGICRGFENSRCYNEM 142
DB 124 DVAEQWKINFAQAJETAKAWTRL 146

RESULT 33
G90107
ubiquitin-conjugating enzyme E2-21 kb [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 15-Jun-2001
C:Accession: G90107
R:Goudas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil
Nature 410, 109-109b, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A9082; MIMB:11323671; PMID:11323671
A:Accession: G90107
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <DXG>
A:Cross-references: GR:AJ010592; NID:ql12580693; PIDN:CAC27017.1; GSPDB:GN00151
C:Genetics:
A:Map position: 2
A:Gene: nucleomorph
C:Superfamily: human ubiquitin-protein ligase E2
C:Keywords: nucleomorph

Query Match 10.3%; Score 136; DB 2; Length 144;
Best Local Similarity 29.4%; Pred. No. 9.5e-05;
Matches 35; Conservative 21; Mismatches 41; Indels 22; Gaps 6;

QY 12 IMVKTFEDRMDFSAALIKGPTRTPEYDGLYLFDIQLNPYVAPVPHFCYLSQCSGRLENPN 71
DB 22 INIKYEDLYKKWGEIIGNGTPYGGKSENFECVPLSVPLSPPKITFVDGI---FHPN 78
QY 72 LP-ENCKVCVSLA-GTWIGKGTERTWTSKSSLLQVLISIQGLIIV-----NEPYNEAG 122
DB 79 VYPSNGRCLDLEKNOW -TPAWT-----ILFSQQAIVLLTNPNFNSPINCADAG 126

RESULT 34
S32672
ubiquitin-protein ligase (EC 6.3.2.19) UBC10 - Arabidopsis thaliana
N:Alternate names: ubiquitin-conjugating enzyme (UBC10)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 03-Jun-2002
C:Accession: S32672
R:Giroud, P.; Carpenter, L.H.; van Nocker, S.; Sullivan, M.L.; Vierstra, R.D.
submitted to the EMBL Data Library, August 1992

A:Description: Homologs of the essential ubiquitin conjugating enzymes ubc1, 4, and
A:Reference number: S32672
A:Accession: S32672
A:Molecule type: mRNA
A:Residues: 1-148 <GIR>
A:Cross-references: EMBL:214991; NID:g297877; PIDN:CAA78715.1; PID:g297878
C:Superfamily: human ubiquitin-protein ligase E2
C:Keywords: ligase

Query Match 10.3%; Score 135; DB 2; Length 148;
Best Local Similarity 30.9%; Pred. No. 0.00012;
Matches 35; Conservative 20; Mismatches 35; Indels 12; Gaps 3;

QY 22 DLF--SALIKGPTRTPEYDGLYLFDIQLNPYVAPVPHFCYLSQCSGRLENLYDNGKVC 79
DB 29 DMFWQATWMSSESPYAGGVFLVIHFPPDYFEPKPKVAFRIKV---FHPNINSNGSIC 85
QY 80 VSLGIGWIGKGTERTWTSKSSLLQVLISIQGLIIVNEP 116
DB 86 LDFI-----KEQSPALTIKSLVLSICSULTPEPNP 115

RESULT 35
T08465
ubiquitin-protein ligase (EC 6.3.2.19) - fruit fly (Drosophila melanogaster)
N:Alternate names: ubiquitin conjugating enzyme
C:Species: Drosophila melanogaster
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
C:Accession: T08465
R:Kirby, R.J.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z16421
A:Accession: T08465
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-199 <KIR>
A:Cross-references: EMBL:X92838; NID:gl359613; PID:gl054722
A:Experimental source: strain Canton-S
C:Genetics:
A:Gene: UbcD4
A:Cross-references: FlyBase:FBgn0015321
C:Superfamily: yeast ubiquitin-protein ligase UBC1
C:Keywords: cell cycle control; DNA repair; ligase; protein degradation

Query Match 10.3%; Score 135; DB 2; Length 199;
Best Local Similarity 23.4%; Pred. No. 0.00018;
Matches 44; Conservative 26; Mismatches 82; Indels 36; Gaps 6;

QY 12 IMVKTFEDRMDFSAALIKGPTRTPEYDGLYLFDIQLNPYVAPVPHFCYLSQCSGRLENPN 71
DB 27 IKIELVNDSWTELGEIAGPPDIPEGGKFKVLEIKVPETYPNPKAREITRI---WHPN 83
QY 72 LYD-NKVCVSLGWIGKGTERTWTSKSSLLQVLISIQGLIIVNEPYNEAGFDSRGGLQ 130
DB 84 ISSVTGATCGLDIL-----KDNMAAAMILRTVLSIQALLAAAE-----DQPDQAV 129
QY 131 EGYENSRKYNEMALIRVQSMITQVRRPPEVF-----EQEIRQHIFSTGG 174
DB 130 VAYOFKDKYDLFELL--TAKHWNAYAGGPHTFPDCDSKIQRUDMGIDHEARAVLSKEN 187
QY 175 WRLVNRIE 182
DB 188 WMLEKATE 195

RESULT 36
A40797
ubiquitin-conjugating enzyme - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 16-Jul-1999
C:Accession: A40797
R:Chen, Z.; Niles, E.G.; Pickart, C.M.
J. Biol. Chem. 266, 15698-15704, 1991

C:Genetics:
A:Gene: FlyBase:eff
A:Cross-references: FlyBase:FBgn0011217
C:Superfamily: human ubiquitin-protein ligase E2

Query Match 10.2% Score 134; DB 2; Length 147;
Best Local Similarity 27.9%; Pred. No. 0.00015;
Matches 36; Conservative 24; Mismatches 43; Indels 26; Gaps 6;

QY 22 DLF--SALIKGTRTPYEGLYFDIOLPNITYPAVPHPFCYLSCSGRLNPNIYDNGKVC 79
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 DLFWHQATIMCPDSYPQGVVFTHFFPDYDFPKPKVAFIRI---YHPNINSNGSIC 85
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 80 VSLGIWTWIGKTRWTSSKLQLVLISIQGLIL---VNEPYYNAG--PDSRCLQGVE 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 LDIL-----RSQSPALTISKVLLSICSLLCDPNDDPLYPEAIRIYKTDR----- 131
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 135 NSRCYNEMA 143
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 --EKYNEIA 138
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 39
S32674
ubiquitin-protein ligase (EC 6.3.2.19) UBC9 - Arabidopsis thaliana
N:Alternate names: protein T13J8.70; ubiquitin-conjugating enzyme UBC9
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 03-Jun-2002
C:Accession: S32674; J02897
R:Cloud, P.; Carpenter, T.B.; van Nocker, S.; Sullivan, M.L.; Vierstra, R.D.
submitted to the EMBL Data Library, August 1992
A:Description: Homologs of the essential ubiquitin conjugating enzymes ubcl, 4,
A:Reference number: S32672
A:Accession: S32674
A:Molecule type: mRNA
A:Residues: 1-148 <GIR>
A:Cross-references: EMBL:Z14990; NID:g297883; PIDN:CAA78714.1; PID:g297884
R:Revan, M.; Pohl, T.; Weizengger, T.; Hoheisel, J.; Meves, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z14766
A:Accession: J02897
A:Molecule type: DNA
A:Residues: 1-148 <BEV>
A:Cross-references: EMBL:A1035524
A:Experimental source: cultivar Columbia; BAC clone T13J8
C:Genetics:
A:Gene: UBC9
A:Map position: 4
A:Intons: 247; 66/3; 101/3
A>Note: T13J8.70
C:Superfamily: human ubiquitin-protein ligase E2
C:Keywords: ligase

Query Match 10.2% Score 134; DB 2; Length 148;
Best Local Similarity 30.9%; Pred. No. 0.00015;
Matches 30; Conservative 20; Mismatches 35; Indels 12; Gaps 3;

QY 22 DLF--SALIKGTRTPYEGLYFDIOLPNITYPAVPHPFCYLSCSGRLNPNIYDNGKVC 79
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 DMFHWQATIMCPDSYPQGVVFTHFFPDYDFPKPKVAFIRTKV---PHNINSNGSIC 85
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 80 VSLGIWTWIGKTRWTSSKLQLVLISIQGLILVNRP 116
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 LDIL-----RKQSPALTISKVLLSICSLLTDNP 115
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 40
S43782
ubiquitin-conjugating enzyme UBC3 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
C:Accession: S43782
R:Sullivan, M.I.; Carpenter, T.B.; Vierstra, R.D.

Plant Mol. Biol. 24, 651-661, 1994
A: Title: Homologues of wheat ubiquitin-conjugating enzymes - TaUBC1 and TaUBC4 are encoded by two genes
A: Reference number: S43781; M010:54207190; PMID:8155884
A: Accession: S43782
A: Status: nucleic acid sequence not shown
A: Molecule type: DNA
A: Residues: 1-150 <SOL>
A: Cross-references: EMBL:U19352; NID:g431261; PID:AAA32896.1; PID:g431262
A: Genes: UBC3
A: Introns: 12/2; 51/1; 81/1; 110/3
C: Superfamily: human ubiquitin-protein ligase E2

Query Match: 10.28; Score 134; DB 2; Length 150;
Best local Similarity 27.68; Pred. No. 9.00015;
Matches 40; Conservative 24; Mismatches 61; Indels 20; Gaps 4;

QY 9 PEGIMKTFEDRMDFSLMIGPIRTPYEDGLYLFDTQLPNIYFVAVPPHFCYLSQCSGRL 68
DB 21 PVGISAPQGNIMHNNALIFGPDTPWDGGTFKLTFTEDYENKPPVIVRSR---MF 77
QY 59 MPNIYENGKVCYSILGQWIGKGTERTWSSQLQVLIISIQGLIIVNPPYNEACFUSDRG 128
DB 78 HPNLYAGSGICLDIL-----QKNSPIYGVAAVLTISQSLCDPNP-----DSPAN 123
QY 129 LQES---YENSRCYNEMALIEVWS 150
DB 124 AFAARLSEKKNKYNKKVIEIVCS 143

Search completed: April 10, 2003, 10:33:41
Job Line : 17.6863 secs

GenCore version 5.1.4.p5.4576
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw mode.

Run on: April 10, 2003, 10:27:50 ; Search time 10.5882 Seconds
(without alignments)
979.302 Million cell updates/sec

Title: US-09-930-026-1

Perfected score: 1315

Sequence: MALLATSLPGRWVWKEPR.....PGEASGDSSESGAAGLAFS 250

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 90

Maximum Match 100

Listing first 135 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	370	28.1	4829	1	R166_HUMAN
2	269	20.5	423	1	YW22_CAEEL
3	179	13.6	172	1	UBC2_YEAST
4	169	12.9	225	1	UBC2_HUMAN
5	165.5	12.6	165	1	UBC7_YEAST
6	161	12.2	153	1	UBC2_YEAST
7	159	12.1	151	1	UBC2_SCHPO
8	153	12.1	152	1	UBC1_CAEEL
9	157	11.9	152	1	UBC2_HUMAN
10	156.5	11.9	166	1	UBC7_SCHPO
11	155.5	11.8	238	1	UBC2_RABIT
12	153	11.6	152	1	UBC2_HUMAN
13	152.5	11.6	151	1	UBC2_NEUCR
14	151	11.5	152	1	UBC2_WHEAT
15	151	11.5	213	1	UBC2_ASEFM
16	150	11.4	179	1	UBC2_CANAL
17	148.5	11.3	152	1	UBC2_MEDSA
18	148	11.3	151	1	UBC6_DROME
19	148	11.3	236	1	UBC3_HUMAN
20	144.5	11.0	152	1	UBC1_ARATH
21	143.5	10.9	152	1	UBC2_ARATH
22	140	10.6	215	1	UBC_ASEFB
23	138.5	10.5	160	1	UBC2_SCHPO
24	138	10.5	148	1	UBC4_LYCES
25	137.5	10.5	145	1	UBC3_HUMAN
26	137	10.4	152	1	UBC2_HUMAN
27	135	10.3	121	1	UBC2_ARATH
28	135	10.3	148	1	UBC6_ARATH
29	134.5	10.3	148	1	UBC4_ARATH
30	134.5	10.2	200	1	UBC1_HUMAN
31	134.5	10.2	204	1	UBC2_PICHA
32	134	10.2	147	1	UBC1_DROME
33	134	10.2	148	1	UBC2_ARATH

34	134	10.2	150	1	UBC3_ARATH
35	134	10.2	199	1	UBC4_DROME
36	133	10.1	147	1	UBC2_CAEEL
37	132.5	10.1	148	1	UBC6_ARATH
38	130	9.9	147	1	UB5B_HUMAN
39	130	9.9	147	1	UB5C_HUMAN
40	129	9.8	218	1	UBC6_YEAST
41	128.5	9.8	177	1	UBC8_SPIPO
42	128.5	9.8	188	1	UBC4_SCHPO
43	128	9.7	147	1	UBC4_SCHPO
44	127	9.7	147	1	UBC4_YEAST
45	127	9.7	157	1	UBC9_YEAST
46	126.5	9.6	164	1	UBC7_CAEEL
47	126.5	9.6	167	1	UBC7_ARATH
48	126.5	9.6	260	1	UBC8_CAEEL
49	126	9.6	151	1	UBC3_DROME
50	125	9.5	147	1	UB5D_RAT
51	124	9.4	147	1	UBC4_CANAL
52	122	9.3	148	1	UBC5_YEAST
53	121	9.2	147	1	UB5A_HUMAN
54	119	9.0	157	1	UBC3_SCHPO
55	119	9.0	250	1	UBC6_YEAST
56	118	9.0	158	1	UBC1_HUMAN
57	117	8.9	170	1	UBC6_HUMAN
58	116.5	8.9	179	1	UBC6_HUMAN
59	115.5	8.8	156	1	UBC4_YEAST
60	115.5	8.8	215	1	UBC1_YEAST
61	115	8.7	158	1	UBC1_MESAU
62	113	8.6	179	1	UBC4_XENLA
63	109	8.3	183	1	UBC4_HUMAN
64	109	8.3	295	1	UBC3_YEAST
65	108	8.2	183	1	UBC3_YEAST
66	107.5	8.2	153	1	UBC7_DROME
67	107.5	8.2	154	1	UBC7_DROME
68	107	8.1	168	1	UBC7_WHEAT
69	105.5	8.0	183	1	UBC6_ARATH
70	101	7.7	176	1	UBC2_SCHPO
71	98	7.6	161	1	UBC2_DROME
72	96.5	7.5	232	1	UBC2_DROME
73	98	7.5	187	1	UBC4_ARATH
74	98	7.5	655	1	UBC1_HUMAN
75	96	7.3	152	1	UBC8_HUMAN
76	95.5	7.3	207	1	UBC6_MOUSE
77	94	7.1	430	1	UBC1_HUMAN
78	92.5	7.0	185	1	UBC5_ARATH
79	92.5	7.0	193	1	UBC6_HUMAN
80	92.5	7.0	193	1	UBC6_MOUSE
81	92.5	7.0	430	1	UBC1_MOUSE
82	91	6.9	2128	1	UBC6_MOUSE
83	87	6.6	3866	1	UBC6_MOUSE
84	86.5	6.6	251	1	KDCA_VIRCH
85	86.5	6.6	384	1	UBC1_HUMAN
86	85	6.5	395	1	UBC1_HUMAN
87	85	6.5	519	1	UBC1_HUMAN
88	83.5	6.3	613	1	UBC1_HUMAN
89	82.5	6.3	550	1	SYR_MYCIN
90	82.5	6.3	1435	1	UBC1_HUMAN
91	82	6.2	459	1	UBC1_HUMAN
92	81	6.2	617	1	UBC1_MOUSE
93	81	6.2	619	1	UBC1_MOUSE
94	81	6.2	2331	1	UBC1_MOUSE
95	80.5	6.1	268	1	UBC1_MOUSE
96	80.5	6.1	417	1	UBC1_MOUSE
97	80.5	6.1	1363	1	UBC1_MOUSE
98	80.5	6.1	1514	1	UBC1_MOUSE
99	80	6.1	550	1	UBC1_MOUSE
100	80	6.1	629	1	UBC1_MOUSE
101	80	6.1	2137	1	UBC1_MOUSE
102	79.5	6.0	139	1	UBC1_MOUSE
103	79.5	6.0	1271	1	UBC1_MOUSE
104	79	6.0	184	1	UBC1_MOUSE
105	79	6.0	554	1	UBC1_MOUSE
106	78.5	6.0	245	1	UBC1_MOUSE

```

107 78.5 6.0 428 1 DEX2_HUMAN
108 78.5 6.0 418 1 UBP8_HUMAN
109 78.5 6.0 458 1 GIL2_HUMAN
110 78.5 6.0 242 1 RPL1_PABVP
111 77.5 5.9 188 1 UBR1_YEAST
112 77.5 5.9 500 1 DP2_RAT
113 77.5 5.9 546 1 SRP8_MOUSE
114 77.5 5.9 1343 1 CAL1_HUMAN
115 77.5 5.9 1597 1 YUS6_HUMAN
116 77.5 5.9 2331 1 RRP2_MABVP
117 77 5.9 417 1 VGID_HSVBS
118 77 5.9 851 1 DYN1_RAT
119 77 5.9 864 1 DYN1_HUMAN
120 77 5.9 1149 1 DRS_MOUSE
121 77 5.9 1729 1 LARP_HUMAN
122 76.5 5.8 184 1 UBC4_WHEAT
123 76.5 5.8 204 1 LAF4_MOUSE
124 76.5 5.8 259 1 RPO5_VASCC
125 76.5 5.8 259 1 RPO5_VASCC
126 76.5 5.8 336 1 LVC2_DETRA
127 76.5 5.8 490 1 E42_MOUSE
128 76.5 5.8 999 1 ORP_CRIGR
129 76 5.8 229 1 VV_P14HA
130 76 5.8 399 1 RRP_P14HA
131 76 5.8 405 1 CAB_ECOLI
132 76 5.8 417 1 PHEB_RAT
133 76 5.8 441 1 SCAB_RAHIT
134 76 5.8 1149 1 Y6G_CAFEL
135 75.5 5.7 214 1 SH3H_MOUSE

```

ALIGNMENTS

```

RESULT 1
BIR6_HUMAN
ID BIR6_HUMAN STANDARD; PRT: 4829 AA.
AC Q9NR09; Q9UL21;
DI 16-OCT-2001 (Rel. 40, Created)
DT 15-OCT-2002 (Rel. 41, Last annotation update)
DE Baculoviral IAP repeat-containing protein 6 (Ubiquitin-conjugating
EN BIR domain enzyme apollon).
GN BIR6 OR KIAAL269.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
PI 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE-20012759; PubMed-10544019;
RA Chen Z., Naito M., Hori S., Mashima T., Yamori T., Tsumoto I.;
RT "A human IAP-family gene, apollon, expressed in human brain cancer
cells."
RL Biochem. Biophys. Res. Commun. 264:847-854(1999).
RN 121
RP SEQUENCE OF 3238-4829 FROM N.A.
RC MEDLINE-20039619; PubMed-10574462;
RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro."
RL DNA Res. 6:337-345(1999).
CC - TISSUE SPECIFICITY: MAY PROTECT CELLS FROM UNDERGOING APOPTOSIS.
CC - SIMILARITY: EXPRESSED IN BRAIN CANCER CELLS.
CC - SIMILARITY: CONTAINS 1 BIR REPEAT.
CC - SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE UBIQUITIN-
CONJUGATING ENZYME FAMILY.
CC This SWISS PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF265555; AAF75772.1;
CC EMBL: AB033115; BAA86603.1;
CC HSSP: Q13490; LOBH.
CC Genew: HGNC:13516; BIRC6.
CC MIM: 605638;
CC InterPro: IPR001370; BIR.
CC InterPro: IPR000608; UBQ_conjugat.
CC Pfam: PF00179; UBQ_con; 1.
CC Pfam: PF0653; BIR; 1.
CC ProDom: PD000461; UBQ_conjugat; 1.
CC SMART: SM00238; BIR; 1.
CC SMART: SM00212; UBQ; 1.
CC PROSITE: PS01282; BIR_REPEAT_1; FALSE_NEG.
CC PROSITE: PS0143; BIR_REPEAT_2; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; FALSE_NEG.
CC PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
CC K0 Apoptosis; Thiol protease inhibitor; Ub1 conjugation pathway; Ligase.
CC FT REPEAT 236 330 BIR.
CC FT DOMAIN 4348 4676 UBIQUITIN-CONJUGATING.
CC FT BINDING 4597 4597 UBIQUITIN (BY SIMILARITY).
CC FT DOMAIN 2 8 POLY-ALA.
CC FT DOMAIN 1632 1640 POLY-ALA.
CC SQ SEQUENCE 4829 AA; 527604 MW; C67126A672D3653 CRC64;

```

Query Match 28.1%; Score 370; DB 1; Length 4829;
Best Local Similarity 40.3%; Pred. No. 5,10-23;
Matches 77; Conservative 36; Mismatches 68; Indels 10; Gaps 5;

```

QY 4 LATSIP----ESIMVKTEDRMDFLSALIKGTRTPYEDGLYLFDTQLNINYPVPPHFC 59
DB LSTSLPLSSSSSVFRCDEBLDMKVLTPADTPANGCFEDVTFQDYESSPP-LV 4615
QY 60 YLSQCSG---RLNPNLYDNGKVCVSLZGTWIGKGTERTSK-SLLQLVLTISQGLIYNE 115
DB 4616 NLETTGSHSVRFENPNLYDNGKVCVSLZGTWIGKGTERTSK-SLLQLVLTISQGLIYNE 4675
QY 116 PYNMGFSDSRGLQGYENSRCYN-EMALIRVQSMTCIVRRPPEVFECEIRQHFSTGG 174
DB 4676 PYNMGFSDSRGLQGYENSRCYN-EMALIRVQSMTCIVRRPPEVFECEIRQHFSTGG 174
QY 175 WRLVNRIESWL 185
DB 4736 VEIMAGCEEWI 4746

```

```

RESULT 2
YWV2_CAEEL
ID YWV2_CAEEL STANDARD; PRT: 423 AA.
AC Q11076;
DI 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 48.1 kDa protein B0403.2 in chromosome X.
GN B0403.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID:6239;
PI 1;
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2.
RA Geisel C.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: SOME, IN THE N-TERMINUS TO UBIQUITIN-CONJUGATING
CC ENZYMES.
CC -----

```


23 I LADPDGKIKVFPNEEDLTDLQVTTGGPECTPYAGGLFRNKKLILCKDFSPAPKGVFLTK 82

64 CSGRLPNLDYNGKVCYSLLTGWIGXGTERMTSKSSLLQVLTISIOGLILVNPYYNEAF 123

83 I---FHPNVGANGECVNVL-----KRDVTAELGIRHVLTLKCLLIHNP---FSAL 229

124 DSDRG---IQGVENSRCYNMALLRVQSMTQLVRRPVEVFQEIQHMFSTGGWNLVNR 181

130 NEEAGRLLENVEE---YAARAKL-----LTEI-----HGGAGG----- 160

162 ESMLETHALLEKAALPNGVYPKAPFAELSPGQEPEDGGPAPGCAEGSGSDSE 241

161 -----PSGRAEAGRIASGTTEASSIDPCAP-----CGP----- 188

242 GGAOG 246

189 GGAEG 193

RESULT 5

ID	UBC7_YEAST	STANDARD	PRT: 165 AA.
AC	Q02159		
DT	01-JUL-1993 (Rel. 26, Created)		
DT	01-JUL-1993 (Rel. 26, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Ubiquitin-conjugating enzyme E2-18 kDa (EC 6.3.2.19)		
DE	(Ubiquitin-protein ligase)		
GN	UBC7 OR QR18 OR YMR022W OR YMR9711.12.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
OX	NCBI_TaxID=4932;		
RI	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=AB320;		
RX	MEDLINE=93003327; PubMed=1327148;		
RT	Vassil A., Boulet A., Decoster F., Fayc G.;		
RT	"QRI8, a novel ubiquitin-conjugating enzyme in Saccharomyces		
RT	cerevisiae.";		
RT	Biochim. Biophys. Acta 1132:211-213(1992).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	MEDLINE=93149278; PubMed=8381213;		
RT	Jungmann J., Reins H.-A., Schobert C., Jentsch S.;		
RT	"Resistance to cadmium mediated by ubiquitin-dependent proteolysis.";		
RT	Nature 361:369-371(1993).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=S288c / AH972;		
RT	Lye G., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;		
RT	Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	X-RAY CRYSTALLOGRAPHY (2.93 ANGSTROMS).		
RC	MEDLINE=97200716; PubMed=9048545;		
RT	Cook W.J., Martin P.D., Edwards B.F.P., Yamazaki K.K., Chau V.;		
RT	"Crystal structure of a class I ubiquitin conjugating enzyme (Ubc7)		
RT	from Saccharomyces cerevisiae at 2.9-A resolution.";		
RT	Biochemistry 36:1621-1627(1997).		
CC	[-]: FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER		
CC	PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED AND		
CC	ABNORMAL PROTEINS. INVOLVED IN RESISTANCE TO CADMIUM		
CC	POISONING.		
CC	[-]: CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +		
CC	diphosphate + protein N-ubiquityllysine.		
CC	[-]: PATHWAY: Ubiquitin conjugation; second step.		
CC	[-]: INDUCTION: BY CADMIUM.		
CC	[-]: MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR		
CC	UBIQUITIN-THIOLESTER FORMATION.		
CC	[-]: SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		

CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X66829; CAA47302.1; -
DR EMBL: X69100; CAA48846.1; -
DR EMBL: X49211; CAA89125.1; -
DR PIR: S28951; S28951.
DR PIR: S29338; S29338.
DR PIR: S29741; S29741.
DR PDB: 2OC2; 1B-MAX-98.
DR SGD: S0004624; GR18.
DR InterPro: IPR000608; UBC_conjugat.
DR Pfam: PF00179; UQ_con: 1.
DR ProDom: PD000461; UBC_conjugat: 1.
DR SMART: SM00212; UBC: 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubi conjugation pathway; Ligase; Multigene family; 3D-structure;
KW Cadmium.
FT RNDING 89 89 UBIQUITIN.
SQ SEQUENCE 165 AA; 18526 MW; D3D2378472BB462D CRC64;

Query Match 12.6%; Score 165.5; DB 1; Length 165;
Best Local Similarity 31.2%; Pred. No. 2.4e-07;
Matches 39; Conservative 28; Mismatches 45; Indels 13; Gaps 5;

QY 9 PEGIMV-KTFEDRMDFSLALIKGTTRPYDGLYLFDIOLPNIYPAVPHPCYLSQCSGR 67
DB 21 PPGIAGKPSSENNIFIDKLLIGPPDPYADGVFNAKLEFPKDPSPKLTFTPSI--- 77

QY 68 LNPALYDNGKVCVSLGLI-----WICKGIERWTSSKSLQVLISIOGLILVNEPYVNEA 121
DB 78 LHPNIYPNGEVCISLHSGDOPNMYELAEKRSVPQSVKEKILLSVMSL--SEPNI-ES 134

QY 122 GFUSD 126
DB 135 GANID 135

RESULT 6
UBC2_YEAST
ID UBC2_YEAST STANDARD; PRT; 153 AA.
AC P52490;
DI 01-OCT-1996 (Rel. 34; Created)
DI 01-OCT-1996 (Rel. 34; Last sequence update)
DI 15-JUN-2002 (Rel. 41; Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17.5 kDa (EC 6.3.2.19) (Ubiquitin-
DE protein ligase) (Ubiquitin carrier protein)
GN UBC13 OR YDR092W OR YD6652.04.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S286c / AB972;
RX MEDLINE-96162026; PubMed-8576256;
RA Matuschewski K., Hauser H.P., Treier M., Jentsch S.;
RT "Identification of a novel family of ubiquitin-conjugating enzymes
RT with distinct amino-terminal extensions.";
RJ J. Biol. Chem. 271:2789-2794(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S286c / AB972;
RA Oliver K., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
RI Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
CC diphosphate + protein N-ubiquityllysine.

CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC STRONGEST, TO DROSOPHILA BEN/UBCD3 AND TO UBC5.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X99443; CAA67806.1; -
DR EMBL: Z50111; CAA90451.1; -
DR HSP: P15731; IQCO.
DR SGD: S00C2459; UBC13.
DR InterPro: IPR000608; UBC_conjugat.
DR Pfam: PF00179; UQ_con: 1.
DR ProDom: PD000461; UBC_conjugat: 1.
DR SMART: SM00212; UBC: 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubi conjugation pathway; Ligase; Multigene family.
FT BINDING 87 87 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 153 AA; 17468 MW; 445558F8F193275B CRC64;

Query Match 12.2%; Score 161; DB 1; Length 153;
Best Local Similarity 29.6%; Pred. No. 5.4e-07;
Matches 37; Conservative 29; Mismatches 45; Indels 14; Gaps 4;

QY 3 LIALSLPEGIMVTFEDRMDFSLALIKGTTRPYDGLYLFDIOLPNIYPAVPHPCYLS 62
DB 15 LVSDPVP-CITAPHDNDLNKYFQVIEGSPSPEDGIFELLYLPDDYPMAPKVRPLT 73

QY 63 QCSGRNPNIYDNGKVCVSLGLIOWTIGKTERWTSSKSLQVLISIOGLIL---VNEPYN 119
DB 74 KI---YHINIDRLGRLGCLDVLKT-----NWSPALQIRTVLLSIGALLASPNPDPLAN 123

QY 120 EAGFD 124
DB 124 DVAED 128

RESULT 7
UBC2_SCHPO
ID UBC2_SCHPO STANDARD; PRT; 151 AA.
AC P23566;
DI 01-NOV-1991 (Rel. 20; Created)
DI 01-NOV-1995 (Rel. 32; Last sequence update)
DI 15-JUN-2002 (Rel. 41; Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19) (Ubiquitin-
DE protein ligase 2) (Ubiquitin carrier protein) (RAD6 homolog).
GN RHP6 OR SPAC18B11.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID:4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90228339; PubMed-2184030;
RA Reynolds P., Koken M.H.M., Hoeijmakers J.H.J., Prakash L.;
RT "The rhp6- gene of Schizosaccharomyces pombe: a structural and
RT functional homolog of the RAD6 gene from the distantly related yeast
RT EMBO J. 9:1423-1430(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RX MEDLINE-21848401; PubMed-11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

```

Qy 129 LQSYVNSRCY 139
  |  |  |
Db 127 AQLHREKKRY 137

RESULT 8
UNCL_CABEI
ID UNCL_CABEI STANDARD; PRT: 192 AA.
AC P52478; O45062;
DC 01-OCT-1996 (Rel. 34, Created);
DT 01-OCT-1996 (Rel. 34, Last sequence update);
DI 15-JUN-2002 (Rel. 41, Last annotation update);
DE Ubiquitin-conjugating enzyme E2-21.5 kDa (EC 6.3.2.19) (Ubiquitin-
protein ligase) (ubiquitin carrier protein);
DN UBC-1 OR C35B1.1.
GC Caenorhabditis elegans.
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodietidae; Caenorhabditis.
OX NCBI_TaxID:6239;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=96027757; PubMed=7546294;
RT Leggett D.S., Jones D., Caeddo E.P.M.;
RT *Caenorhabditis elegans UBC-1, a ubiquitin-conjugating enzyme
RT homologous to yeast RAB6/UBC, contains a novel carboxy-terminal
RT extension that is conserved in nematodes.;
RL DNA Cell Biol. 14:883-891(1995).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Miller N., Stelljes L., Bradshaw H., Keppier D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
REVISIONS.
RA Waterston R.;
RC Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CL -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC STRONGEST, TO YEAST UBC2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

QY 9 PEGIMVKTEEDRMUFLSALIKGPTPTDYDGLVLDIQLPNIYPAVPHFYLSQCSGRL 68
 DB 21 PAVGSCAPTEEDNITLWEATTFSPGETPFDDTFTKLSLSETEHYPNKPPTVKFISK---MF 77
 QY 69 NPNIYNGKVCVSLIGIWI:GAGTIRWTSSKSLQVLIS:QGLILVNGPYNYEAGFOSDRG 128
 DB 78 HPNVIYAGSGICIDIL-----QNRWSPYVVAIIISIQSL--DEPNNSPA--NSLA 126
 QY 129 LQGYENSRNYEMALIRVQSMIO 153
 DB 127 AQ:YQENRREYK-----NVOQIVHQ 147
 RESULT 9
 UBC2_HUMAN
 TU UBC2_HUMAN STANDARD; PRT; 152 AA.
 AC P21567;
 DT 01-NOV-1991 (Rel. 20, Created)
 DI 01-NOV-1991 (Rel. 20, Last sequence update)
 DI 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2 B (EC 6.3.2.19) (Ubiquitin-protein
 ligase B) (Ubiquitin carrier protein B) (HR6B) (HHR6B) (E2-17 kDa).
 GN UBE2B OR RAB6B
 OS Homo sapiens (Human).
 OS Mus musculus (Mouse).
 OS Rattus norvegicus (Rat), and
 CS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID:9606, 10090, 10116, 9986;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC SPECIES-Human;
 RX MEDLINE:90228340; PubMed:2158443;
 RA Schneider R., Ekerskorn C., Lottspeich F., Schweiger M.;
 RT "The human ubiquitin carrier protein E2(Mr ~ 17,000) is homologous to
 RL the yeast DNA repair gene RAD6.";
 RL EMBO J. 9:1431-1435(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human;
 RX MEDLINE:92020951; PubMed:1717990;
 RA Koken M.H.M., Reynolds P., Jaspers Dekker I., Prakash L., Prakash S.,
 RA Bootsma D., Hoeijmakers J.H.J.;
 RT "Structural and functional conservation of two human homologs of the
 RT yeast DNA repair gene RAD6.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8865-8869(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human;
 RA Poloumienko A., Hlecher S.R.;
 RT "Exon-intron structure of mammalian HR6A and HR6B genes.";
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBS databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Rabbit, and Rat;
 RX MEDLINE:92202189; PubMed:1313008;
 RA Wang S.S., Gumas F., Hanville D.;
 RT "A rabbit reticulocyte ubiquitin carrier protein that supports
 RT ubiquitin dependent proteolysis (E2:4k) is homologous to the yeast
 RT DNA repair gene RAD6.";
 RL J. Biol. Chem. 267:6495-6501(1992).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Rat;
 RX MEDLINE:94324482; PubMed:8048511;
 RA Wang S.S., Banville D.;
 RT "14-kDa ubiquitin-conjugating enzyme: structure of the rat gene and
 RL regulation upon fasting and by insulin.";
 RL Am. J. Physiol. 267:E339-E348(1994).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Mouse; STRAIN=C57BL/6 X CBA; Tissue=Testis;

RA Roest H.P., van Klaveren J., de Wit J., van Gurp C.G., Koken M.H.M.,
 RA Vermeij M., van Rooijen J.H., Vreeburg J.T.M., Baarends W.M.,
 RA Bootsma D., Grooten J.A., Hoeijmakers J.H.J.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBS databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Mouse;
 RA Varshavsky A., Griqoryev S., Stewart A.F., Kwon Y.T., Arfin S.M.,
 RA Bradshaw R.A., Jenkins N.A., Copeland N.G.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBS databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. REQUIRED FOR POSTREPLICATION REPAIR OF UV-DAMAGED
 CC DNA.
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- SUBUNIT: Interacts with RAD18.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC STRONGEST TO YEAST UBC2
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M74525; AAA35982.1;
 DR EMBL: X53251; AAA37339.1;
 DR EMBL: AF294392; AAG27628.1;
 DR EMBL: AF294387; AAG27628.1; JOINED.
 DR EMBL: AF294388; AAG27628.1; JOINED.
 DR EMBL: AF294389; AAG27628.1; JOINED.
 DR EMBL: AF294390; AAG27628.1; JOINED.
 DR EMBL: AF294391; AAG27628.1; JOINED.
 DR EMBL: M62387; AAG31492.1;
 DR EMBL: M62388; AAG21087.1;
 DR EMBL: U04308; AAB60669.1;
 DR EMBL: U04303; AAB60669.1; JOINED.
 DR EMBL: U04304; AAB60669.1; JOINED.
 DR EMBL: U04305; AAB60669.1; JOINED.
 DR EMBL: U04306; AAB60669.1; JOINED.
 DR EMBL: U04307; AAB60669.1; JOINED.
 DR EMBL: X96859; CAA65802.1;
 DR EMBL: U57690; AAC52884.1;
 DR PIR: S12530; S12530.
 DR PIR: B41222; B41222.
 DR PIR: A42416; A42416.
 DR HSSP: P25865; 2AAK.
 DR Genew: HNCN:12473; UBE2B.
 DR MIN: 179095.
 DR MGD: MGI:102944; Ube2b.
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF00179; UO_con; 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBCc; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ubl conjugation pathway; Ligase; DNA repair; Nuclear protein;
 KW Multigene family.
 FT: BINDING 88 88 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 152 AA; 17312 MW; 6FDDEE7E06840BF CRC64;
 Query Match 11.9%; Score 157; DB 1; Length 152;
 Best Local Similarity 30.3%; Pred. NC. 1.2e-06;
 Matches 43; Conservative 25; Mismatches 60; Indels 14; Gaps 4;
 OY 9 PEGIMVKTEEDRMUFLSALIKGPTPTDYDGLVLDIQLPNIYPAVPHFYLSQCSGRL 68
 DB 21 PAVGSCAPSENNIMONNAVIFGPESTPFDDTFTKLSLSETEHYPNKPPTVKFISK---MF 77

[illegible]

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

DR EMBL: U58652; AAB02655.1; ALT_INIT.
 DR HSSP: Q02159; 2UC2.
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: PD000461; UQ_conjugat; 1.
 DR SMART: SM00212; UBC2; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW UBI conjugation pathway: Ligase; Multigene family.
 F: BINDING 93 UBIQUITIN (BY SIMILARITY).
 FT DOMAIN 260 238 ASP/GLU-RICH (ACIDIC).
 SQ SEQUENCE 238 AA; 17166 MW; 1896CF0116A56308 CRC64;

Query Match 11.8%; Score 155.5; DB 1; Length 238;
 Best Local Similarity 32.8%; Pred. No. 2.8e-06;
 Matches 41; Conservative 25; Mismatches 44; Indels 15; Gaps 6;

OY 10 FGIMVKTFEDRMOLFSAALGKPIPTPYEDGLYLPDIQIPNLYPAVPPHFCYLSQCSGR 67
 Db 26 EEFKI-TLVDSGLYNVEALFPPNTLYEGYFAKHKFPIDYKSPPTFRFLK---M 81
 OY 68 LNPILYDNGKVCVSLGTSWICK-----GTERWTSKSLIQVLISQGLVNEP-YNNE 120
 Db 82 WHEINIVNGVCISLHPPVDMQSGELPSRWNPQINVRTLLISV---ISLNEPNTSP 139
 OY 121 ASFDS 125
 Db 140 ANVDA 144

RESULT 12
 UBC2_HUMAN
 ID UBC2_HUMAN STANDARD; PRT; 152 AA.
 AC M9459;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2 A (EC 6.3.2.19) (Ubiquitin-protein
 DE ligase A) (Ubiquitin carrier protein A) (HR6A) (HR6A).
 GN UBE2A OR KADEA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92020951; PubMed=177690;
 RA Koken M.H.M., Reynolds P., Jaspers-Dekker I., Prakash L., Prakash S.,
 RA Boelsma E., Boelmakers J.H.J.;
 RT Structural and functional conservation of two human homologs of the
 RT yeast DNA repair gene RAD6.*;
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8865-8869;1991;
 EN 12.

RP SEQUENCE FROM N.A.
 RA Polomanko A., Biecher S.R.;
 RT Exon-intron structure of mammalian UBE2A and UBE2B genes.*;
 RL Submitted (AUG 2000) to the EMBL/Genbank/DBJ databases.
 CC 1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC DNA.
 CC OTHER PROTEINS REQUIRED FOR POSTREPLICATION REPAIR OF UV-DAMAGED
 CC DNA.

CC 1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC 1- PATHWAY: Ubiquitin conjugation; second step.
 CC 1- SUBUNIT: Interacts with RAD18.
 CC 1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOESTER FORMATION.
 CC 1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC STRONGEST, TO YEAST UBC2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

DR EMBL: M74524; AAA35981.1; --
 DR EMBL: AF294366; AAG27430.1; --
 DR EMBL: AF294361; AAG27430.1; JOINED.
 DR EMBL: AF294362; AAG27430.1; JOINED.
 DR EMBL: AF294363; AAG27430.1; JOINED.
 DR EMBL: AF294364; AAG27430.1; JOINED.
 DR EMBL: AF294365; AAG27430.1; JOINED.
 DR HSSP: P25865; 2AAK.
 DR Genew: HGNC:12472; UBE2A.
 DR MIM: 312180; --
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: PD000461; UQ_conjugat; 1.
 DR SMART: SM00212; UBC2; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW UBI conjugation pathway: Ligase; DNA repair; Nuclear protein;
 KW Multigene family.
 FT BINDING 88 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 152 AA; 17243 MW; 7AB6173D5FA6DE1 CRC64;

Query Match 11.6%; Score 153; DB 1; Length 152;
 Best Local Similarity 28.9%; Pred. No. 2.6e-06;
 Matches 41; Conservative 27; Mismatches 60; Indels 14; Gaps 4;

OY 9 PRCIMVKTFFEDRMOLFSAALGKPIPTPYEDGLYLPDIQIPNLYPAVPPHFCYLSQCSGR 68
 Db 21 PAGVSGAPSENNIMVNAVIFGPEPTFGDGTFLKLTIEFTTEYPNKPPTVRFYSK---MF 77
 OY 69 NPNLYDNGKVCVSLGTSWICK-----GTERWTSKSLIQVLISQGLVNEP-YNNEAGFSDHC 128
 Db 78 HPNVYADGSGICIDIL-----QNRWSPDYVSSILTSIOSLL--DEPNPSA--NSQA 126
 OY 129 IQEGYENSCYNEMALIRVQS 150
 Db 127 AOLYQENKREYKRVSAIVFQS 148

RESULT 13
 UBC2_NEUCR
 ID UBC2_NEUCR STANDARD; PRT; 151 AA.
 AC P25453;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
 DE (Ubiquitin-protein ligase 2) (Ubiquitin carrier protein).
 GN MUS-8.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96337904; PubMed=8753651;
 RA Sosh T., Sakuraba Y., Kafer E., Inoue H.;
 RT The mus-8 gene of Neurospora crassa encodes a structural and
 RT functional homolog of the Rad6 protein of Saccharomyces cerevisiae.*;
 RL Curr. Genet. 30:224-231;1996).
 CC 1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC 1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC 1- PATHWAY: Ubiquitin conjugation; second step.

```

CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC STRONGEST, TO YEAST UBC2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D78372; RAAL1380.1; -
CC HSSP: P06104; IAYZ.
CC InterPro: IPR000608; UBQ_conjugat.
CC Pfam: PF00179; UQ_con; 1.
CC ProDom: PD000461; UBQ_conjugat; 1.
CC SMART: SM00212; URCG; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
CC Ubl conjugation pathway; Ligase; Multigene family.
CC BINDING 88 88 UBIQUITIN.
CC MUTAGEN 88 88 C->S: LOSS OF ACTIVITY.
CC SEQUENCE 151 AA: 17204 MW: 551401458CB6DC81 CRC64;
CC -----
CC Query Match 11.6%; Score 152.5; DB 1; Length 151;
CC Best Local Similarity 27.5%; Pred. No. 2.8e-06;
CC Matches 38; Conservative 24; Mismatches 55; Indels 21; Gaps 4;
CC -----
CC 4 LATSLEPIGVKTFEDRMDFSLALIKGPIRTPYEDGLVLEDIQIPNLYPAVPPHPCYLSQ 63
CC Db 16 MAFEDPACVASPVENWVNNVAVIGPADTFEDGIFRLVHIFEDYKPKPSWKFISE 75
CC QY 64 CSGRINPNLYKNGKVCVSLGWTWIKGTERTWSKSSLOVLSTIQGLILVNPYYNEAGF 123
CC Db 76 ---MEHPNVAATGECEDIL-----QNRKSPYDVAALVLSIQSL--NSP----- 116
CC QY 124 DSDRLGEGYENSKYNE 143
CC Db 117 ---NTGSPANVEASNLKMD 132
CC -----
CC RESULT 14
CC UBC2_WHEAT
CC ID UBC2_WHEAT STANDARD; PRI: 152 AA.
CC AC P25865;
CC DT 01-MAY-1992 (Rel. 22, Created)
CC DT 01-MAY-1992 (Rel. 22, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Ubiqutin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
CC DE (Ubiqutin-protein ligase) (Ubiqutin carrier protein).
CC GN UBC1.
CC OS Triticum aestivum (Wheat).
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC OC Poaceae; Triticum.
CC OX NCBI_TaxID=4565;
CC [1]
CC SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-88.
CC STRAIN-cv. Augusta;
CC RC MEDLINE-92084684; PubMed-1560887;
CC RA Sullivan M.L., Viersira R.D.;
CC "Cloning of a 16-kDa ubiquitin carrier protein from wheat and
CC Arabidopsis thaliana. Identification of functional domains by in
CC vitro mutagenesis."
CC RL J. Biol. Chem. 266:23878-23885(1991);
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC OTHER PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -----

```

```

CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M62720; AAA34310.1; -
CC HSSP: P25865; 2AAK.
CC InterPro: IPR000608; UBQ_conjugat.
CC Pfam: PF00179; UQ_con; 1.
CC ProDom: PD000461; UBQ_conjugat; 1.
CC SMART: SM00212; URCG; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
CC Ubl conjugation pathway; Ligase; Multigene family.
CC BINDING 88 88 UBIQUITIN.
CC MUTAGEN 88 88 C->S: LOSS OF ACTIVITY.
CC SEQUENCE 152 AA: 17301 MW: 513AA5439F84C8AH CRC64;
CC -----
CC Query Match 11.5%; Score 151; DB 1; Length 152;
CC Best Local Similarity 28.6%; Pred. No. 3.8e-06;
CC Matches 42; Conservative 26; Mismatches 59; Indels 20; Gaps 4;
CC -----
CC 9 PEGIMVKTTFEDRMDFSLALIKGPIRTPYEDGLVLEDIQIPNLYPAVPPHPCYLSQ 68
CC Db 21 PAGISAPHDNNITLWNNVIFGDDIPMDGGIFKTLQFTEDYKPKPTVREYSR---MF 77
CC QY 69 NPNLYDNGKVCVSLGWTWIKGTERTWSKSSLOVLSTIQGLILVNPYYNEAGDS 125
CC Db 78 HPNLYAGSICLDIL-----QNRKSPYDVAALVLSIQSLCPNPNPANEAA--- 127
CC QY 126 DRGLQEGYENSKYNEALIRVVSMT 152
CC Db 128 ----RMYSENKREYNRKRVREVSMT 150
CC -----
CC RESULT 15
CC UBC2_ASF2
CC ID UBC2_ASF2 STANDARD; PRI: 213 AA.
CC AC P25869;
CC DT 01-MAY-1992 (Rel. 22, Created)
CC DT 01-MAY-1992 (Rel. 22, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Ubiqutin-conjugating enzyme E2-21 kDa (EC 6.3.2.19)
CC DE (Ubiqutin-protein ligase) (Ubiqutin carrier protein).
CC GN ASFV-UBC.
CC OS African swine fever virus (isolate Malawi Lil 20/1) (ASFV).
CC OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
CC OX NCBI_TaxID=10500;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE-92155177; PubMed-1310934;
CC RA Hingamp P.M., Arnold J.E., Mayer R.J., Dixon L.K.;
CC "A ubiquitin conjugating enzyme encoded by African swine fever
CC virus."
CC RL EMBO J. 11:361-366(1992).
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC OTHER PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



```

DR EMBL: L06967; AAA18526.1;
DR HSSP: P25855; 2AAK.
DR InterPro: IPR000638; UBQ_conjugat.
DR Pfam: PF00179; UBQ_con; 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBQ; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS50127; UBIQUITIN_CONJUGAT_2; 1.
DR Ubl conjugation pathway; Ligase; Multigene family.
FT BINDING 88 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 152 AA: 17315 MW: 3951040CF5C6E24 CRC64;

Query Match 11.3%; Score 148.5; DB 1; Length 152;
Best Local Similarity 27.2%; Pred. No. 6.3e-06;
Matches 44; Conservative 26; Mismatches 54; Indels 35; Gaps 5;

QY 9 PEGIVKVFTEPDMDFSLNLIKGRTPTRPEDGGLYFDIQLPNITYPAPVPHFCYLSCSGRL 68
DB 21 PAVICSGAPQDNIMWNAVIFQPDTPWCGCFK:SLQFSDYPNKPPIVRVSR---MP 77
CY 69 NPNLYNKKVSVLLGTWIGKTERWTSKSLQLVLSIQGLII---VNEPYNEAGFDS 125
DB 78 HPNVADGGICLDIL-----QNWSPDYDVAALTSIQSLCDPNPNSPANSEAA--- 127
QY 126 DRGICGAGYNSKCYNEALIRVQVS:QIVKRPPEVFQ 164
DB 128 ----RMFSNKREYN-----RRVREWFQ 147

RESULT 18
UBC6_DROME STANDARD; PRT: 151 AA.
AC P25153; U9VNT70;
DT 01-MAY-1992 (Rel. 22, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiqutin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
DE (Ubiqutin-protein ligase) (Ubiqutin carrier protein).
GN UBC6 OR DHR6 OR CG2013.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
RX MEDLINE-91219466; PubMed-1902572;
RA Koker M.H.M., Reynolds P., Bootsma D., Hoeijmakers J.H.J., Prakash S.,
RA Prakash S.;
RA "Dhr6, a Drosophila homolog of the yeast DNA-repair gene RAD6.";
RA Proc. Natl. Acad. Sci. U.S.A. 88:3832-3836(1991).
[2]
SEQUENCE FROM N.A.
RX STRAIN Berkeley;
RX MEDLINE-23196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfliffer B.D.,
RA Wan K.H., Doyle C., Baxter H.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Bersch K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Brock J., Brookstein P., Brotlier P.,
RA Hurtis K.C., Husam D.A., Hutler H., Cadieu F., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.H., Davies P.,
RA de Pablo H., Delcher A., Deng Z., Mays A.P., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Duqar-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garq N.S., Geibart W.M., Glasser K.,
RA Glöck A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D., Heiman T.J., Hernandez C.R., Hottel J.,

```

```

RA Hostie D., Houston K.A., Howland J.J., Moi M.-H., Ibeqwan C.,
RA Ja'ali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relnert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith I.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA *The genome sequence of Drosophila melanogaster.*;
RA Science 287:2185-2195(2000).
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC DNA.
CC -!- OTHER PROTEINS. REQUIRED FOR POSTREPLICATION REPAIR OF UV-DAMAGED
CC DNA.
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: Ubiquitin conjugation; second stop.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC STRONGEST, TO YEAST RAD6.

```

```

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)

```

```

EMBL: M63792; AAA28308.1;
EMBL: M64435; AAA28309.1;
EMBL: M63791; AAA28309.1; JOINED.
EMBL: AE003604; AAF52079.1;
HSSP: A39392; A39392.
FlyBase: FBgn0004436; UbcD6.
InterPro: IPR000608; UBQ_conjugat.
Pfam: PF00179; UBQ_con; 1.
ProDom: PD000461; UBQ_conjugat; 1.
SMART: SM00212; UBQ; 1.
PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
PROSITE: PS50127; UBIQUITIN_CONJUGAT_2; 1.
Ubl conjugation pathway; Ligase; DNA repair; Nuclear protein;
Multigene family.
FT BINDING 88 88 UBIQUITIN (BY SIMILARITY).
FT CONFLICT 99 99 T -> R (IN REF. 1).
SQ SEQUENCE 151 AA: 17152 MW: CC4B35992E4A9220 CRC64;

```

```

Query Match 11.3%; Score 148; DB 1; Length 151;
Best Local Similarity 28.2%; Pred. No. 6.9e-06;
Matches 40; Conservative 29; Mismatches 59; Indels 14; Gaps 4;

QY 9 PEGIVKVFTEPDMDFSLNLIKGRTPTRPEDGGLYFDIQLPNITYPAPVPHFCYLSCSGRL 68
DB 21 PAVICSGAPQDNIMWNAVIFQPDTPWCGCFK:TLFFTEEPNKPPIVRVSR---F 77
QY 69 NPNLYNKKVSVLLGTWIGKTERWTSKSLQLVLSIQGLII---VNEPYNEAGFDSRG 128
DB 78 HPNVADGGICLDIL-----QNRWSPDYDVAALTSIQSLCDPNPNSPANSTAAQ 128
QY 129 LOEGYNSRCYNEMALIRVQVS 150

```



```

Db 125 LYK -ENBREYKRVKACVUS 148
RESULT 19
UHC3 HUMAN
ID UHC3 HUMAN STANDARD: PRT: 236 AA.
AC P49427;
DE 01-FEB-1996 (Rel. 33, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-32 kDa complementing (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (E2-CDC34).
DE CDC34.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID:9606;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE-94068425; PubMed-8248134;
RA Pion S.E., Jeppig K.A., De H.N., Groudine M.;
RT "Cloning of the human homolog of the CDC34 cell cycle gene by
RT complementation in yeast."
RL Proc. Natl. Acad. Sci. U.S.A. 90:10484-10488(1993).
RN 12;
RP SEQUENCE FROM N.A.
RC 1-SSHEL-brain, and Lung;
RA Strausberg R.;
RL Submitted (FEF-2002) to the EMBL/GenBank/CDL databases.
CC 1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS.
CC 1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
CC diphosphate + protein N-ubiquityl-lysine.
CC 1- PATHWAY: Ubiquitin conjugation; second step.
CC 1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC 1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC STRONGEST, TO YEAST UBC3.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L22005; AAC37534.1; ALI_INT.
CC EMBL: BC009850; AAR09850.1; -.
CC EMBL: BC024643; NAH18143.1; -.
CC EMBL: BC023979; AAR23979.1; -.
CC HSSP: Q02159; 20CZ.
CC Genew: HGNC:1734; CDC34.
CC MIM: 116948; -.
CC InterPro: IPR006068; Ubq_conjugat.
CC Pfam: PF00179; Ubq_conj; -.
CC ProDom: PD000461; Ubq_conjugat; 1.
CC SMART: SM00212; UBC3; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
CC UBL conjugation pathway; Ligase; Multigene family.
FT BINDING 53 93 UBIQUITIN (BY SIMILARITY).
FT DOMAIN 200 236 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 236 AA; 26737 MW; 25896066B589DB3 CRC64;
Query Match 11.38; Score 148; DB 1; Length 236;
Best Local Similarity 30.38; Pred. No. 1.2e-05;
Matches 44; Conservative 28; Mismatches 55; Indels 18; Gaps 7;
OY 10 EGIWVIFERMDLFS--ALIKGPRTPYEDGLYIFDIQLNIVPAVPHFCVLSOCGR 67
1: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 26 EGRV-ILVDEGDLNWEVAIFGPNTPYEGGYFKARCKFPIDYPSPPAFLIK---M 81
1: | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

OY 58 LNPXLYDNGKVCVSLIGTWIGK-----GTERWTSKSLQLVLSIQGLILVNEF--YNE 120
1: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 82 WHPNIYETGVCIILRPDPVDPQSGELPSPWRNPTONVNTILSV--ISLNEPTFSP 139
1: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 121 AGFSD--RGLOPYENSRCYNEM 142
1: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 140 ANVDASVMYRKWKESGKDEYIDI 164
1: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 20
UBCLARATH
ID UBCLARATH STANDARD: PRT: 152 AA.
AC P25855;
DE 01-MAY-1992 (Rel. 22, Created)
DE 01-MAY-1992 (Rel. 22, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 1 (EC 6.3.2.19) (Ubiquitin-
DE protein ligase 1) (Ubiquitin carrier protein 1).
DE UBCL OR AT1G14400 OR F14L17_35 OR F14L17_17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN 12;
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RX MEDLINE-92084684; PubMed-1660887;
RA Sullivan M.L., Vierstra R.D.;
RT "Cloning of a 16-kDa ubiquitin carrier protein from wheat and
RT Arabidopsis thaliana. Identification of functional domains by in
RT vitro mutagenesis."
RL J. Biol. Chem. 266:23878-23885(1991).
RN 12;
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RX MEDLINE-94207190; PubMed-8155884;
RA Sullivan M.L., Carpenter T.B., Vierstra R.D.;
RT "Homologues of wheat ubiquitin-conjugating enzymes -- TaUCL1 and
RT TaUBC4 are encoded by small multigene families in Arabidopsis
RL plant Mol. Biol. 24:651-661(1994).
RN 3;
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RX MEDLINE-21016719; PubMed-11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buchler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gili J.E., Golusmith A.D., Haas R., Hanson N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jonkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.-S., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maltli R., Marzilli A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Saizberg S.L., Schwartz J.R., Shim P., Southwick A.M.,
RA Sun B., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Ueberback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
RN 4;
RP SEQUENCE OF 1-98 FROM N.A.
RC STRAIN-cv. Columbia;
RA Raynal M., Greilet F., Iaudic M., Moyer Y., Cooke R., Deiseny M.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN 5;
RX X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE-92340563; PubMed-1321826;

```



```

CC -----
DR EMBL: L19353; AAA32899.1; -
DR EMBL: Y13031; CAA73476.1; -
DR EMBL: AC202521; AAC05346.1; -
DR HSSP: P25865; 2NAK.
DR InterPro: IPR000608; UQ_conjugat.
DR ProDom: PD000461; UQ_conjugat; 1.
DR SMART: SM0212; UQC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubl conjugation pathway; Ligase; Multigene family.
FT BINDING 88 88 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 152 AA; 17279 MW; 266CD41930C38629 CRC64;

Query Match 10.9%; Score 143.5; DB 1; Length 152;
Best Local Similarity 26.4%; Pred. No. 1.7e-05;
Matches 42; Conservative 29; Mismatches 53; Indels 35; Gaps 5;

QY 9 PGIMVKTFFDMFLPSALIKGPTPTPEYDGLYLFDIQLPNIPVPPHFCYLSQCSGR 68
DB 21 PAIGSAPQNNIMLWNAVIGPDTPWDGGTFKZSLQFSESDYKPKPTVRFVSR---MF 77
QY 69 NPLYDNGKVCVSLIG:WICKGTHFWTSKSSLLQVLSISLQGLIL---VNPYPYNEAGFDS 125
DB 78 HPN:YAKS:CLDIL-----GNQSPYCYVAAITSLQSLCDPN:NPANSEA---- 126
QY 126 DRGLOEYENSKYCNEMALIRVVQSMTCIVRRPPEVFQ 164
DB 127 ---AHMT-SF-----SKREYNRRVRVVFQ 147

RESULT 22
UBCG_ASFB7
ID UBCG_ASFB7 STANDARD: PRT; 215 AA.
AC P27949;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-21 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
CN [215];
OS African swine fever virus (Strain BAFV) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID:10498;
RN [1]
RX MEDLINE:92087485; PubMed:1309282;
RA Rodriguez J.M., Salas M.L., Vinuela E.;
RT "Genes homologous to ubiquitin-conjugating proteins and eukaryotic
transcription factor SII in African swine fever virus.";
RL Virology 186:40-52(1992).
RN [2]
RX COMPLETE GENOME.
RA Vanez R.J., Rodriguez J.M., Nosal M.L., Yuste L., Enriquez C.,
Rodriguez J.F., Vinuela E.;
RT "Analysis of the complete nucleotide sequence of African swine fever
virus.";
RL Virology 208:249-278(1995).
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
OTHER PROTEINS.
CC -!- CATALYTIC ACTIVITY: AIP + ubiquitin + protein lysine -> AMP +
diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC -!- DEVELOPMENTAL STAGE: MAINLY LATE IN INFECTION.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
UBIQUITIN THIOLESTER FORMATION.
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: M77121; AAA42704.1; -
DR EMBL: U18466; AAA65370.1; -
DR PIR: F39448; OXFAS.
DR HSSP: Q02159; 2UCZ.
DR InterPro: IPR000608; UQ_conjugat.
DR ProDom: PD000461; UQ_conjugat; 1.
DR SMART: SM0212; UQC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubl conjugation pathway; Ligase; Late protein.
FT BINDING 185 85 UBIQUITIN (BY SIMILARITY).
FT DOMAIN 183 215 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 215 AA; 24737 MW; 3488R6C87E67727 CRC64;

Query Match 10.6%; Score 140; DB 1; Length 215;
Best Local Similarity 27.9%; Pred. No. 5.3e-05;
Matches 31; Conservative 25; Mismatches 43; Indels 12; Gaps 4;

QY 10 PGIMVKTFFDMFLPSALIKGPTPTPEYDGLYLFDIQLPNIPVPPHFCYLSQCSGR 69
DB 19 ENFKISVNNENITENDVILRGPPDTLYEGCLFKAKVAFPPPYAPPKLFTSE---MMH 75
QY 70 NPLYDNGKVCVSLIGTWIGKTER----WTSKSSLLQVLSISLQGLILVNEP 116
DB 76 PNYPGRICTISLH---GDNARFGMTWSPAQKIDITILLSV--ISLNEP 121

RESULT 23
UBCG_SCHPO
ID UBCG_SCHPO STANDARD: PRT; 160 AA.
AC Q9P611;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
CN UBC16 OR SPBC1198.09.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae.
OX NCBI_TaxID:4896;
RN [1]
RX SEQUENCE FROM N.A.
RA STRAIN=972;
RA MEDLINE:21848401; PubMed:11859160;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
Squires J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
Olliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
Waltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
Gabel C., Fuchs M., Fritz C., Holzer F., Moestl D., Hilbert H.,
Horzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
Gallbert F., Aves S.J., Xiao Z., Hunt C., Moore K., Mottier S.M.,
Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Henito J.,
Domigues A., Revuelta J., Moreno S., Armstrong J., Forsburg S.L.,

```


RA Shintani A., Sasaki T., Nagamine K., Mitsuoyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordstok G., Hornschler K., Brandt P.,
 RA Scherf M., Schoen A., Desario A., Reichelt J., Kauer G., Boecker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann E., Dagand E.,
 RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lebrach H., Reinhardt R., Yaspo M.-L.,
 RT *cDNA sequence of human chromosome 21.*;
 PL Nature 405:311-319(2000).
 RN [3]
 RF SEQUENCE FROM N.A.
 RC SPECIES-Human; TISSUE-Lucy;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBS databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Mouse; STRAIN-C57BL/6J; TISSUE-Fetal;
 RX MEDLINE-21238254; PubMed-11278356;
 RA Tiwari S., Weissman A.M.;
 RT "Endoplasmic reticulum (ER)-associated degradation of 1 cell receptor
 subunits. Involvement of ER-associated ubiquitin-conjugating enzymes
 (E2s).";
 RL J. Biol. Chem. 276:16193-16200(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Mouse;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBS databases.
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 OTHER PROTEINS.
 CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine ~ AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -1- PATHWAY: Ubiquitin conjugation; second step.
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC STRONGEST, TO C.ELEGANS UBC7.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF032456; AAC32312.1;
 CC EMBL: AL163300; CAB00551.1;
 CC EMBL: BC001736; AAH01736.1;
 CC EMBL: BC008351; AAH08351.1;
 CC EMBL: BC011569; AAH11569.1;
 CC EMBL: AF296657; AAK52608.1;
 CC EMBL: BC013321; AAK10321.1;
 CC HSP: 202159; 28CZ
 CC Genew: HGNC:112483; J02262.
 CC MIM: 603124;
 CC MIM: 603124;
 CC InterPro: IPR000608; UBC_conjugat.
 CC Pfam: PF00179; UQ_con; 1.
 CC Prodom: P0000461; UBC_conjugat; 1.
 CC SMART: SM00212; UBC; 1.
 CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 CC PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 CC Ubl conjugation pathway; Ligase; Multigene family.
 KW UBIQUITIN (BY SIMILARITY).
 FT BINDING 89 89
 FT CONFID 12 12 E -> V (IN REF. 1).
 FT CONFID 101 107 MYESSA -> HGLRFQ (IN REF. 1).
 SQ SEQUENCE 165 AA; 18566 MW; 74DFC732A79575E3 CRC64;
 Query Match 10.58; Score 137.5; DB 1; Length 165;
 Best Local Similarity 29.28; Pred. No. 6.2e-05;
 Matches 38; Conservative 27; Mismatches 52; Indels 13; Gaps 5;
 QY 4 LAISLPEG:WVKIF-EDRMDF:LSALIKGPURIPYEGGLY:EDIQ:PNLY:AVPHPCYLS 62
 DB 15 LAISLPEG:WVKIF-EDRMDF:LSALIKGPURIPYEGGLY:EDIQ:PNLY:AVPHPCYLS 73

Db 16 LTLNPEGVAGPNEENFENFALLMGPDTCFVGFAILSFPLDYPLSPKMRPTC 75
 QY 63 QCSGRLENLYDNCKVCVSLGT-----WKGKIEWTKSSLLQVLISIGLLVNEP 116
 Db 76 E---MFHNIPVPGRCVTSILHAPGDOPMGVSSAERMSVQSVKILLSVSMIA--EP 130
 QY 117 YYNFAGFSD 126
 Db 131 -NDESGANDV 139
 RESULT 26
 UBCN_HUMAN STANDARD; PRT; 152 AA.
 AC Q16781.
 DT 01-NOV-1997 (Rel. 35, Created)
 DI 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2 N (EC 6.3.2.19) (Ubiquitin-protein:
 DE ligase N) (Ubiquitin carrier protein N) (UBC13).
 GN UBE2N.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_taxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97058291; PubMed-8902611;
 RA Yanaguchi T., Kim N.-S., Sekine H., Osaka F., Yamao F.,
 RA Kato S.;
 RT "Cloning and expression of cDNA encoding a human ubiquitin-conjugating
 RT enzyme similar to the Drosophila bendless gene product.";
 RL J. Biochem. 120:494-497(1996).
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS.
 CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine ~ AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -1- PATHWAY: Ubiquitin conjugation; second step.
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION.
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC STRONGEST, TO DROSOPHILA BEN/UBCD3 AND YEAST UBC13.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D83004; BAA11675.1;
 CC HSSP: P15731; LOCC.
 CC Genew: HGNC:12492; UBE2N.
 CC MIM: 603679;
 CC InterPro: IPR000608; UBC_conjugat.
 CC Pfam: PF00179; UQ_con; 1.
 CC Prodom: P0000461; UBC_conjugat; 1.
 CC SMART: SM00212; UBC; 1.
 CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 CC PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
 CC Ubl conjugation pathway; Ligase; Multigene family.
 KW UBIQUITIN (BY SIMILARITY).
 FT BINDING 87 87
 SQ SEQUENCE 152 AA; 17138 MW; FACD84D883D77407 CRC64;
 Query Match 10.48; Score 137; DB 1; Length 152;
 Best Local Similarity 24.58; Pred. No. 6.1e-05;
 Matches 35; Conservative 31; Mismatches 63; Indels 14; Gaps 4;
 QY 3 LLATSLPEGIWVTFEDRMDF:LSALIKGPTRTPVEDGLVLFQIQLPNLY:PAVPHPCYLS 62
 DB 15 LLAPVVP-GIAKAPDPSNARFYHVVVIAGPQDSFPGTFLKLEFLPERYPMAKPVRFMI 73

```

QY 63 QCSGRNLNPNLNDKGVSVLIGTWKSGTERWISKSSLIQVLSIQGIIV---NPPYNN 119
DB 74 K1 --YHPNVKLGRLGRLDLS-----KDKWSPALQRTVLLSIALSAPNPDPLAN 123
QY 120 RAGFDSURGURGSYNSCYNEM 142
DB 124 DVALQWKTNEACAIHTARAWIKC 146

RESULT 27
UNCC_ARATH STANDARD; PRF; 121 AA.
AC P36617;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 12 (EC 6.3.2.19) (Ubiquitin-
protein ligase 12) (Ubiquitin carrier protein 12) (Fragment).
GN UBC12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Leaf;
RX MEDLINE=94035158; PubMed=8220461;
RA Gird P.-A., Carpenter I.B., van Nocker S., Sullivan M.L.,
RA Vierstra R.D.;
RT "Homologs of the essential ubiquitin conjugating enzymes UBC1, 4, and
RT 5 in yeast are encoded by a multigene family in Arabidopsis
thaliana."
RL Plant J. 3:545-552(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=97471969; PubMed=5330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT Pi clones."
RL DNA Res. 4:215-230(1997).
RN [3]
RP SEQUENCE OF 1-74 FROM N.A.
RC STRAIN=cv. Columbia;
RA Bardet C., Dabos P., Tremoussaye D., Lescure B.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 36-148 FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Green silicles;
RA Raynal M., Grillet F., Laidie M., Meyer Y., Cooke R., Delseny M.;
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED
AND ABNORMAL PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN YOUNG STEMS, OLD
LEAVES. LOWEST LEVELS IN FLORAL BUDS, ANTHERS AND YOUNG LEAVES.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
UBIQUITIN-THIOLESTER FORMATION.
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: Z14993; NOT_ANNOTATED_CDS.
CC HSP: P15731; LOC.
CC InterPro: IPR000608; Ubq_conjugat.
CC Pfam: PF00179; Ubq_con: 1.
CC ProDom: P5600461; Ubq_conjugat; 1.
CC SMART: SM00212; UBCc; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS50127; UBIQUITIN_CONJUGAT_2; 1.
CC Ubq_conjugation pathway; ligase; Multigene family.
CC NON_TER 1
CC TER 1
CC SEQUENCE 121 AA; CPACAG4B54B5897 CRC64;

Query Match: 10.3%; Score 135; DB 1; Length 121;
Best local Similarity 30.9%; Pred. No. 6,8e-05;
Matches 30; Conservative 20; Mismatches 35; Indels 12; Gaps 3;

QY 22 DLF--SALLKGPTRTPEDGSLYLFDLQLNLNYPVPPHFCYLSQCSGRNLNPNLNDKGV 79
DB 10 DMFHQAQIMGPSRSPYAGGVFLVTHRPDPYDFKPKVAFRIKV---FHPNINSNGSIC 66

```

```

QY 80 VSLIGTWKSGTERWISKSSLIQVLSIQGIIVNPP 116
DB 67 LDIL-----KEQWSPALTISKVLSISLITDNP 96

RESULT 28
UNCC_ARATH STANDARD; PRF; 148 AA.
AC P35131; Q42308; Q43276;
DT 01-FEB-1994 (Rel. 26, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 8 (EC 6.3.2.19) (Ubiquitin-
protein ligase 8) (Ubiquitin carrier protein 8) (UBCAT4A).
GN UBC8 OR UBC4A OR A15641700 OR MBK23.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Leaf;
RX MEDLINE=94035158; PubMed=8220461;
RA Gird P.-A., Carpenter T.B., van Nocker S., Sullivan M.L.,
RA Vierstra R.D.;
RT "Homologs of the essential ubiquitin conjugating enzymes UBC1, 4, and
RT 5 in yeast are encoded by a multigene family in Arabidopsis
thaliana."
RL Plant J. 3:545-552(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=97471969; PubMed=5330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT Pi clones."
RL DNA Res. 4:215-230(1997).
RN [3]
RP SEQUENCE OF 1-74 FROM N.A.
RC STRAIN=cv. Columbia;
RA Bardet C., Dabos P., Tremoussaye D., Lescure B.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 36-148 FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Green silicles;
RA Raynal M., Grillet F., Laidie M., Meyer Y., Cooke R., Delseny M.;
RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED
AND ABNORMAL PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN YOUNG STEMS, OLD
LEAVES. LOWEST LEVELS IN FLORAL BUDS, ANTHERS AND YOUNG LEAVES.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
UBIQUITIN-THIOLESTER FORMATION.
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: Z14989; CAA78713.1; -.
CC EMBL: AB005233; BAB11476.1; -.
DR

```

```

DR EMBL: Z37225; CAA95527.1;
DR EMBL: Z17692; CAA79036.1;
DR HSSP: P15731; LOCQ
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UBQ_con; 1.
DR ProDom: PD000467; UBQ_conjugat; 1.
DR SMART: SM00212; UBCG; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubl conjugation pathway; Ligase; Multigene family.
FT BINDING 72 72 K -> M (IN REF. 3).
FT CONFIDICT 60 80 S -> D (IN REF. 4).
FT CONFIDICT 98 98 T -> P (IN REF. 4).
FT CONFIDICT 98 98 T -> P (IN REF. 4).
SQ SEQUENCE 148 AA: 16533 MW: 27CNAENBBE74972 CRC64;

Query Match 10.3%; Score 135; DB 1; Length 148;
Best Local Similarity 30.9%; Pred. No. 8.8e-05;
Matches 30; Conservative 19; Mismatches 36; Indels 12; Gaps 3;

QY 22 DLF--SALIKGPTTPEDGLYLFDIOJNIPYAPVPHFCYLSQCSGRNLNLYDNCKVC 79
DQ 29 DMEHQATIMGPSYAGGVFLVTHPEPPYFPKPKVAFRTKV---FHPNINSNGSIC 85
QY 80 VSLGCTWIGKTERWTSKSSLSQVLSIQGLILVNEP 116
DQ 86 IDLI-----KQWSPALTIISKVLLSTGSLTDPNP 115

RESULT 29
UBCL_HUMAN STANDARD; PRT; 148 AA.
AC P35133;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 18 (EC 5.3.2.19) (Ubiquitin-
protein ligase 10) (Ubiquitin carrier protein 10).
GN UBC1 OR AF5653300 OR K1091.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids 1; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID:3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia; T'SSPE-Leaf;
RX MEDL:94035158; PubMed:8229451;
RA Girard P.-A., Carpenter T.B., van Nocker S., Sullivan M.L.,
RA Vicaterra R.D.;
RT *Homologs of the essential ubiquitin conjugating enzymes UBCL 4, and
RT 5 in yeast are encoded by a multigene family in Arabidopsis
RI thaliana.
RL plant J. 3:545-552(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RX MEDLINE:98403884; PubMed-9734875;
RA Kotani H., Nakamura Y., Sato S., Asanizu E., Kaneko T., Miyajima N.,
RA Jabata S.;
RT *Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones.
RL DNA Res. 5:203-216(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. Columbia;
RA Shinozaki K., Davis R.W., Fekker J.R., Theologis A.;
RT *RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).
RL submitted (DEC 2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED

```

```

CC AND ABNORMAL PROTEINS.
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
CC diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z14991; CAN78715.1;
CC EMBL: L00640; AAA32895.1;
CC EMBL: AB013388; BAB09792.1;
CC EMBL: AF326872; AAG41454.1;
CC PIR: S32672; S32672.
CC HSSP: P15731; LOCQ
CC InterPro: IPR000608; UBQ_conjugat.
CC Pfam: PF00179; UBQ_con; 1.
CC ProDom: PD000461; UBQ_conjugat; 1.
CC SMART: SM00212; UBCG; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ubl conjugation pathway; Ligase; Multigene family.
FT BINDING 85 85
FT BINDING 85 85
FT BINDING 85 85
SQ SEQUENCE 148 AA: 16537 MW: 2B83EDCIADZAK657 CRC64;

Query Match 10.3%; Score 135; DB 1; Length 148;
Best Local Similarity 30.9%; Pred. No. 8.8e-05;
Matches 30; Conservative 20; Mismatches 35; Indels 12; Gaps 3;

QY 22 DLF--SALIKGPTTPEDGLYLFDIOJNIPYAPVPHFCYLSQCSGRNLNLYDNCKVC 79
DQ 29 DMEHQATIMGPSYAGGVFLVTHPEPPYFPKPKVAFRTKV---FHPNINSNGSIC 85
QY 80 VSLGCTWIGKTERWTSKSSLSQVLSIQGLILVNEP 116
DQ 86 IDLI-----KQWSPALTIISKVLLSTGSLTDPNP 115

RESULT 30
UBCL_HUMAN STANDARD; PRT; 200 AA.
AC P27924; Q16721;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-25 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (Huntingtin
interacting protein) (HIP-2).
GN HIP2.
OS Homo sapiens (Human), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606, 9913;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Human;
RX MEDLINE:96325051; PubMed-8702625;
RA Kaichman M.A., Graham R.K., Xia G., Koide H.R., Hodgson J.G.,
RA Graham K.C., Goldberg Y.P., Gietz R.D., Pickart C.M., Hayden M.R.;
RT *Huntingtin is ubiquitinated and interacts with a specific ubiquitin-
conjugating enzyme.
RL J. Biol. Chem. 271:19385-19394(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Bovine;

```

[illegible]


```

DE 155 - - - VLLNDP 161
RESULT 32
UBCL_DROME
ID UBC1_DROME STANDARD: PRI: 147 AA.
AC P25867; C9VFC6;
DT 01-MAY-1992 (Rel. 22, Created)
DI 01-MAY-1992 (Rel. 22, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (E2fe2
DE protein)
DE EFF OR UBCD1 OR C37425.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID:7227;
RN [1]
RP SEQUENCE FROM N.A.
RC S:RA: N Oregon-P2; PubMed:1310935;
RX MHU: NE-92151178;
RT Treier M., Seifert W., Jentsch S.;
RT "Drosophila UbcD1 encodes a highly conserved ubiquitin-conjugating
RT enzyme involved in selective protein degradation."
RT EMBO J. 11:367-372(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Geilker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthakrishnan P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballow K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de la Fabre S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Deep S.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.P., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris M.L., Harvey D., Heiman T.J., Hernandez J.R., Horck J.,
RA Hustlin D., Houston K.A., Howland T.J., Wei M.-H., Iregam C.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimchi B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Jasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklav G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weissbeck G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.-A.,
RA Ye J., Yeh K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng H.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.C.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE OF 111 FROM N.A., AND FUNCTION.
RX MEDLINE-97260538; PubMed-9126638;
RA Cenci G., Rawson R.B., Belloni G., Castiglioni D.H., Tudor M.,

```

```

RA Petrucci R., Goldberg M.L., Wasserman S.A., Gatti M.;
RT "UbcD1, a Drosophila ubiquitin-conjugating enzyme required for proper
RT telomere behavior."
RL Genes Dev. 11:863-875(1997).
CC -i- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED AND
CC ABNORMAL PROTEINS. REQUIRED FOR PROPER TELOMERE BEHAVIOR DURING
CC CELL DIVISIONS AND POSSIBLY FOR UBIQUITINATION OF PROTEINS
CC INVOLVED IN POSTMEIOTIC STAGES OF SPERMATOGENESIS. DELETION
CC MUTATIONS ARE LETHAL IN HOMODIPLOIDS.
CC -i- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine -> AMP +
CC diphosphate -> protein N-ubiquityllysine.
CC -i- PATHWAY: ubiquitin conjugation; second step.
CC -i- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC -i- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; XG2575; CA44453.1; -
DR EMBL; AE003706; AAF55094.1; -
DR EMBL; U68298; AAB39622.1; -
DR PIR; S19157; S19157.
DR HSP; P15731; LQCC.
DR FlyBase; FBgn0011217; eff.
DR InterPro; IPR000508; UBQ_conjugat.
DR Pfam; PF00179; UQ_con; 1.
DR ProDom; PD000461; UBQ_conjugat; 1.
DR SMART; SM00212; UBQC; 1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE; PS00127; UBIQUITIN_CONJUGAT_2; 1.
DR Ubl conjugation pathway; Ligase; Melosis; Cell cycle; Cell division;
KW Multigene family.
FT BINDING 85
SQ SEQUENCE 147 AA; 16678 MW; 82E0CE24F925E2FE CRC64;
Query Match 10.28; Score 134; DB 1; Length 147;
Best Local Similarity 27.98; Pred. No. 0.00011;
Matches 36; Conservative 24; Mismatches 43; Indels 26; Gaps 6;
QY 22 DLF--SALIKGTPRPYEDGLYLFDTQLPNLYPAVPPHPCVLSQCSRLNPNLYDNGKVC 79
DB ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 80 VSLGCTWICKGTERWTSKSLIQVLISIQGLIL--VNHPPYNEAG--FSDRLGLQEGYE 134
DB ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 135 NSRCYNEMA 143
DB |||||
DB 132 --EKYNFLA 138
RESULT 33
UBC9_ARATH
ID UBC9_ARATH STANDARD: PRI: 148 AA.
AC P35132; O42014;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 9 (EC 6.3.2.19) (Ubiquitin-
DE protein ligase 9) (Ubiquitin carrier protein 9) (UBCAT4B).
GN UBC9 OR UBC4B OR A14G27960 OR T1338.70.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

```

CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia; TISSUE=Leaf;
 RX MEDLINE=94035156; PubMed=8220461;
 RA Giron P.-A., Carpenter I.H., van Nocker S., Sullivan M.L.,
 RA Vierstra R.D.;
 RT "Homologs of the essential ubiquitin conjugating enzymes UBC1, 4, and
 RT 5 in yeast are encoded by a multigene family in Arabidopsis
 RT thaliana.";
 RL Plant J. 3:545-552(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia;
 RX MEDLINE=94359481; PubMed=8078492;
 RA Genschik P., Durr A., Fleck J.;
 RT "Differential expression of several E2-type ubiquitin carrier protein
 RT genes at different developmental stages in Arabidopsis thaliana and
 RT Nicotiana glauca.";
 RL Mol. Gen. Syst. 244:548-556(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia;
 RX MEDLINE=26083488; PubMed 16617199;
 RA Mayer K.P.X., Schueller C., Wambitt R., Murphy G., Voelckert G.,
 RA Pohl T., Doesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselmeier M., de Simone V., Obermaier B., Maché R., Mueller T.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert H., Portetelle J., Perez-Alonso M., Boutry M., Hancock I.,
 RA Vos P., Hobeisel J., Zimmermann W., Wedler H., Radley P.,
 RA Landham S.-A., McQuilagh B., William L., Robben J.,
 RA Van der Schueren J., Gryanoprez B., Chiang Y.-J., Vandenbussche F.,
 RA Bracken M., Welljens J., Voet M., Bastiaens I., Aert R., Deloor E.,
 RA Weitzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirksen W.,
 RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Bernieris S., Hoppel S., Feldpausch M., Lambirth S., Van den Daele H.,
 RA De Keyser A., Buysshuert C., Gietlen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Gronin A., Quail M., Bray-Allen S.,
 RA Clark L., Duggett J., Hall S., Kay M., Leonard N., McLay K., Mayes R.,
 RA Pettetti A., Rajadream M.A., Lynne M., Benes V., Kechmann S.,
 RA Horkova D., Hockert B., Scharfe M., Grimm M., Koehnert T.-H.,
 RA Jose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Arqariou A., Vitale D., Liguori R., Piravandi E.,
 RA Massonell O., Oulley F., Cabaud G., Muenzlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Cherdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons I., Weber N., Vandenbol M., Hargues M., Terol J., Torres A.,
 RA Perez-Perez A., Parnelle B., Bent E., Johnson S., Tacon D., Jesse I.,
 RA Heinen L., Schwarz S., Scholler P., Heber S., Francis P., Bielek C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Hovan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj I., Schutz K., Huang E., Spiegele I.,
 RA Sektion M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramar J., Pulten L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hofman J., Tili S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamced A., Iodhi M., Johnson A.,
 RA Chen E., Marra M., Martenssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:769-777(1999).
 RN [4]
 RP SEQUENCE OF 1-82 FROM N.A.
 RC STRAIN-cv. Columbia;
 RA Berthomieu P., Guerrier D., Giraudat J.;
 RL Submitted (NOV-1992) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED
 CC AND ABNORMAL PROTEINS.
 CC -1- CATALYTIC ACTIVITY: Amp - ubiquitin + protein lysine - AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -1- PATHWAY: Ubiquitin conjugation; second step.
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN YOUNG STEMS, OLD
 CC LEAVES, LOWEST LEVELS IN FLORAL BUDS, ANthers AND YOUNG LEAVES.
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
 CC UBIQUITIN-THIOLESTER FORMATION.
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC
 DR EMBL: Z14990; CAA78714.1;
 DR EMBL: X72526; CAA51201.1;
 DR EMBL: L00639; AAA32894.1;
 DR EMBL: AL035524; CAB36765.1;
 DR EMBL: AL161572; CAB79598.1;
 DR EMBL: Z18473; CAA79198.1;
 DR PIR: S12674; S32674.
 DR RSP: P15731; I000.
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: P0000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBCG; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
 KW UBI conjugation pathway; Ligase; Multigene family.
 FT BINDING 85 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 148 AA; 16553 MW; 17F247FC1801BBD CRC64;
 Query Match 10.28; Score 134; DB 1; Length 148;
 Best Local Similarity 30.98; Pred. No. 0.00011;
 Matches 30; Conservative 20; Mismatches 35; Indels 12; Gaps 3;
 QY 22 DLF--SALIKPTPTDGLYDFDQLNPIYPPHPCYLCSCRLNPLYDNGKVC 79
 Db 29 DNEHQWATIMPSDPSYSGGVFLVTIHPDPYFPKPKVAFRTKV---FHPNINSNGSIC 85
 QY 80 VSLICIGKGIHWTSKSLQLQVLSIQGLIIVNEP 116
 Db 86 LDIL-----KEQWSPALITSKVLISICSLITDPNP 115
 RESULT 34
 UBC3_ARATH
 ID UBC3_ARATH STANDARD; PRT; 150 AA.
 AC P42746;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-17 kDa 3 (pI 6.3.2.19) (Ubiquitin-
 DE protein ligase 3) (Ubiquitin carrier protein 3).
 GN UBC3 OR AT5G62540 OR K19815.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia; TISSUE=Green leaf;
 RX MEDLINE=94207190; PubMed=8155884;
 RA Sullivan M.L., Carpenter T.B., Vierstra R.D.;

```

*Homologues of wheat ubiquitin-conjugating enzymes -- TaUBC1 and
TaUBC4 are encoded by small multigene families in Arabidopsis
thaliana.
[2]
Plant Mol. Biol. 24:651-661(1994).
RN NCBI_TaxID=7227;
RN SEQUENCE FROM N.A.
RC STRAIN=cy. Columbia;
EX MEDLINE=99087489; PubMed=9872454;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;
*Structural analysis of Arabidopsis thaliana chromosome 5. VII.
Sequence features of the regions of 1,013,767 bp covered by sixteen
physically assigned P1 and PAC clones.
DNA Res. 5:297-308(1996);
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
OTHER PROTEINS. REQUIRED FOR POSTREPLICATION REPAIR OF UV-DAMAGED
DNA (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED. LOWER
LEVELS FOUND IN LEAVES.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
UBIQUITIN-THIOLESTER FORMATION
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
EMBL: U19352; AAA32898.1;
DR EMBL: AB05469; BAB1504.1;
DR HSP: P25065; 2AAK
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UBQ_con; 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBCG; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW UBI conjugation pathway; ligase; Multigene family.
FT BINDING 88 88 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 150 AA; 17136 MW; 993FE149D541F2AA CRC64;

Query Match 10.2%; Score 134; DB 1; Length 150;
Best Local Similarity 27.6%; Pred. No. 0.00012;
Matches 40; Conservative 24; Mismatches 61; Indels 20; Gaps 4;

QY 9 PPGIMVKTFFDMDFLSALIKGTPRIYPRDGLYFDIQLPNYPAVPPHFCYLSQCSGRU 68
Dh 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
21 PVGISGAPCDNNHWNALIFGPDTPWDSGGTFKLTLFHTEDYENKPIVRSR---MF 77
QY 69 NNVLNDKVCVSLGATGWCKGTWKTSKSLQLVLISQGLILVNEPYNNEAGFDSRG 128
Dh 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
78 HNIYADSGICDIL-----QNWSPYIVAAVLISQSLCDPNP-----DSPAN 123
QY 129 LQEG---YENSRCYNEALIRVQVS 150
Dh 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
124 AFARLFSENKREYNRKVIVEQS 148

RESULT 35
ID UBQ4_DROME STANDARD; PRT: 199 AA.
AC P52486;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-22 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
DE Caenorhabditis elegans.

```

```

GN UBCD4.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN NCBI_TaxID=7227;
RN SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RA Kirby R.J.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
PROTEINS.
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
diphosphate + protein N-ubiquityllysine.
CC -1- PATHWAY: Ubiquitin conjugation; second step.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
UBIQUITIN-THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
STRONGEST, TO YEAST UBC1.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
EMBL: X2838; CAA63424.1;
DR HSP: P15731; IQCQ.
DR FlyBase; FBgn015321; Ubcd4.
DR InterPro: IPR000449; UBA_domain.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UBQ_con; 1.
DR Pfam: PF00627; UBA; 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00165; UBA; 1.
DR SMART: SM00212; UBCG; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW UBI conjugation pathway; ligase; Multigene family.
FT BINDING 92 92 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 199 AA; 22391 MW; 720CA9595FC0C08F CRC64;

Query Match 10.2%; Score 134; DB 1; Length 199;
Best Local Similarity 24.4%; Pred. No. 0.00016;
Matches 42; Conservative 24; Mismatches 70; Indels 36; Gaps 6;

QY 28 IKGPTRIPEYDGLYFDIQLPNYPAVPPHFCYLSQCSGRNPNLYD-NGKVCVSLGTM 86
Dh 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
43 IAGPDDTPVEGKGVLEIKVPTTPYENPKAREIN[---WHPNLSSTVGAICLDIL]--- 96
QY 87 IKGITERWTSKSLQLVLISQGLILVNEPYNNEAGFDSRGLOEGYENSRCYNEALIR 146
Dh 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
97 -----KQWAAATLTLLVLSQALLAAEP-----DDPQAVVAYQFKDYDFLL-- 143
QY 147 VQSMQLVRRPEVF-----FOELRQHFSTGGWRLLVNRIF 182
Dh 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
144 TAKIWTNAYAGPHTFPDCKSTQRLRDMGIDEHARAVLSKENWLEKATE 195

RESULT 36
ID UBQ2_CAEEL STANDARD; PRT: 147 AA.
AC P35129;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).
DE UBC-2 OR LEI-70 OR M7.1.
DE Caenorhabditis elegans.

```

RESULT: 37

UBCH_ARATH
ID UBCH_ARATH STANDARD: PRI: 148 AA.
AC P35134; Q9M9J1;
CT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme F2-17 kDa 11 (EC 6.3.2.19) (Ubiquitin-
protein ligase 11) (Ubiquitin carrier protein 11).
GN UBQ11 OR AT3G08690 OR F17014.16.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLIN=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,
RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delseny M., Boutry M., Grivell L.A., Mache R., Pulgomenech P.,
RA De Simone V., Cholsne N., Artiguenave F., Robert C., Brottier P.,
RA Winkler P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Renes V.,
RA Wiedemann R., Drzonek H., Hoff H., Jordan N., Randert S.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T. H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Hargues M., Ierol J., Clement J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwälder H., Duchemir D.,
RA Cooke R., Laudie M., Berger-Llauró C., Purnelle H., Masny D.,
RA De Haan M., Maarso A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Hewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Otterback T., Fujii C.Y., Shea T.P.,
RA Cready T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Miltischer J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
RI "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RI thaliana.";
RI Nature 408:820-822(2000).
RN [2]
RP SEQUENCE OF 31-148 FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Leaf;
RX MEDLIN=94035158; PubMed=8220461;
RA Giroud P.-A., Carpenter T.B., van Nocker S., Sullivan M.L.,
RA Vicstris R.D.;
RI "Homologs of the essential ubiquitin conjugating enzymes UBC1, 4, and
RI 5 in yeast are encoded by a multigene family in Arabidopsis
RI thaliana.";
RI Plant J. 3:545-552(1993).
RL
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC OTHER PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED
CC AND ABNORMAL PROTEINS.
CC
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine - AMP +
CC diphosphate + protein N-ubiquityllysine.
CC
CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC UBIQUITIN-THIOLESTER FORMATION.
CC
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC purposes is allowed only on request from the EMBL outstation at the
CC European Bioinformatics Institute.

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL: AC012562; AAC51362.1; -
CC EMBL: Z14992; CAA787.6.1; -
CC EMBL: L00641; AAA32896.1; -
CC PIR: S32673; S32673.
CC USSP: P15731; IQCC.
CC InterPro: IPR006038; UBLQ_conjugat.
CC Pfam: PF00179; UQ_conj.
CC ProDom: PD000461; UBLQ_conjugat; 1.
CC SMART: SM00212; UBCC; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_2; 1.
CC UBL conjugation pathway; Ligase; Multigene family.
CC BINDING: 85 85 UBIQUITIN (BY SIMILARITY).
CC SEQUENCE: 148 AA; 16551 MW; 7440992.115BA50EA CRC64;

Query Match
Best Local Similarity 10.1%; Score 132.5; DB 1; Length 148;
Matches 39; Conservative 23; Mismatches 47; Indels 17; Gaps 5;

QY 22 DIF-SALIKGPIRTPYEDGLYFDIQLPNLYPAVPPHFCY;SQCGRINPLNYDNCKVC 79
DB 29 DMHFWQATIMGPSPYAGGVFLVSIHPPDPYKPKYSEKTKV---YHPNINSNGSI 85
QY 80 VSLICLWIGKIGKHWTSKSSLLQVLISLGLIL---VNEPPYNEAG--FDSRGIQEGYE 134
DB 86 LDIL-----KQWSPALITISKVLLSICSLIDPNDPLVPEIAHMYKIDRSKYESTA 138
QY 135 NS 136
DB 139 RS 140

RESULT 38
UB5C_HUMAN STANDARD: PRT: 147 AA.
AC P51669;
D 01-OCT-1996 (Rel. 34, Created)
D 01-OCT-1996 (Rel. 34, Last sequence update)
D 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ubiquitin-conjugating enzyme E2-17 kDa 2 (EC 6.3.2.19) (Ubiquitin-
protein ligase) (Ubiquitin carrier protein) (E2(17)KB 2).
GN UBE2D2 OR UBC2H5 OR UBC4.
OS Homo sapiens (Human).
OS Mus musculus (Mouse).
OS Rattus norvegicus (Rat), and
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606, 10090, 10116, 8355;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Human.
BX MEDLINE-96107191; PubMed-8530467;
KA Jensen J.P., Bates P.W., Tang M., Vierstra R.D., Weissman A.M.;
RT "Identification of a family of closely related human ubiquitin
conjugating enzymes.*"
RL J. Biol. Chem. 270:30408-30414(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Human.
BX MEDLINE-96107191; PubMed-7724550;
KA Rolfe M., Beer-Romero P., Glass S., Eckstein J., Berdo I.,
RA Theodoras A., Pagano M., Draetta G.;
RT "Reconstitution of p53-ubiquitination reactions from purified
components: the role of human ubiquitin-conjugating enzyme UBC4 and
E6 associated protein (E6AP).*"
RL Proc. Natl. Acad. Sci. U.S.A. 92:3264-3268(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-Rat; TISSUE-Testis;

```

```

EX MEDLINE-95126893; PubMed-7826319;
RA Wang S.S., Jain P.;
RT "Molecular cloning, expression and characterization of a ubiquitin
conjugation enzyme (E2(17)KB) highly expressed in rat testis.*"
RL Biochem. J. 305:125-132(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES-Mouse;
RA Richardson P., Zon L.I.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES-X.laevis;
RX MEDLINE-96298869; PubMed-8723350;
RA Yu H., King R.W., Peters J.M., Kirschner M.W.;
RT "Identification of a novel ubiquitin-conjugating enzyme involved in
mitotic cyclin degradation.*"
RL Curr. Biol. 6:455-466(1996).
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
PROTEINS. MEDIATES THE SELECTIVE DEGRADATION OF SHORT-LIVED AND
ABNORMAL PROTEINS. FUNCTIONS IN THE E6/P5-AP-INDUCED
UBIQUITINATION OF P53.
CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine + AMP +
diphosphate + protein N-ubiquityllysine.
CC -!- PATHWAY: Ubiquitin conjugation; second step.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
URTOUITIN-THIOLESTER FORMATION.
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
```

```

CC -----
CC EMBL: U39317; AAA91460.1; -
CC EMBL: L40146; AAC41750.1; -
CC EMBL: U13176; AAA85101.1; -
CC EMBL: U62483; AAR05772.1; -
CC USSP: P15731; IQCC.
CC Genew: HGNC:12475; UBE2D2.
CC MIM: 602962;
CC MGD: MGI:1930715; Ubc2d2
CC InterPro: IPR000608; UBLQ_conjugat.
CC Pfam: PF00179; UQ_conj; 1.
CC ProDom: PD000461; UBLQ_conjugat; 1.
CC SMART: SM00212; UBCC; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
CC PROSITE: PS00183; UBIQUITIN_CONJUGAT_2; 1.
KW Ubl conjugation pathway; Ligase; Multigene family.
FT BINDING: 85 85 UBIQUITIN (BY SIMILARITY).
FT CONFLICT: 128 128 K -> Q (IN REF. 2).
SQ SEQUENCE 147 AA; 16735 MW; C942BE7853CBC355 CRC64;

Query Match
Best Local Similarity 9.9%; Score 130; DB 1; Length 147;
Matches 28; Conservative 20; Mismatches 40; Indels 10; Gaps 2;

QY 19 DRMDLFSALIKGPTPTPYEDGLYFDIQLPNLYPAVPPHFCY;SQCGRINPLNYDNCKVC 78
DB 28 DMHFWQATIMGPNDSPYGGVFLITHTFDYFKPKPKVAFTTTL---YHPNINSNGSI 84
QY 79 CVSLIGTWIGKIGKHWTSKSSLLQVLISLGLILVNEP 116
DB 85 CLDIL-----RSQWSPALITISKVLLSICSLCDPNDP 115

RESULT 39
UB5C_HUMAN STANDARD: PRT: 147 AA.
ID UB5C_HUMAN
AC P47986;

```

```

QY 19 DRMDLFSALINGKTRTPYEGGLYLFDIQLPNIPYAPVPHFCYLSQGRILNPLNDGKV 78
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 DMFPHQATINGNDSPYOGGWFELTHFTPTDPFPKPKVATTRI---YHPNINSNGSI 84
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 79 CYSVLGTTGWKGTERTSKSSLLQVLSIQGLILVNEP 116
      || || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 CLDIL-----RSQSPALITISKVLLSICLLCDPNP 115
      || || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULI 40
UBC8_YEAST
ID   UBC8_YEAST                STANDARD;             PRT;       218 AA.
AC   P28263;
DT   01-DEC-1992 (Rel. 24, Created)
DI   01-JUN-1994 (Rel. 29, Last sequence update)
DE   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Ubiquitin-conjugating enzyme E2-24 kDa (EC 6.3.2.19)
DE   (Ubiquitin-protein ligase) (ubiquitin carrier protein).
GN   UBC8 OR YEL012W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX   NCBI_TaxID:4932;
RN   [1]
RS   SEQUENCE FROM N.A.
RP   MEDLINE-9132089; PubMed-1869573;
RA   Qir S., Nakajima B., Nomura M., Arfin S.M.;
RT   "Cloning and characterization of a Saccharomyces cerevisiae gene
RT   encoding a new member of the ubiquitin-conjugating protein family.";
RL   J. Biol. Chem. 266:15549-15554(1991).
RL   [2]
RS   SEQUENCE FROM N.A.
RP   STRAIN-S288C / AH972;
RA   Dietrich F.S., Mulligan J.T., Hennessey K.M., Ailion E., Araujo R.,
RA   Aviles E., Berino A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA   Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA   Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA   Mossdale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA   Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA   Taylor P., Wei Y., Yelton M., Hotstein D., Davis K.W.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RL   [3]
RN   REVISIONS TO N-TERMINUS.
RP   MEDLINE-94179285; PubMed-8132613;
RA   Karzer P., Seufert W., Hoefler L., Kofler B., Sachsenmaier C.,
RA   Herzig P., Jenisch S., Schweiger M., Schneider R.;
RT   "A human ubiquitin-conjugating enzyme homologous to yeast UBC8.";
RL   J. Biol. Chem. 269:8797-8802(1994).
CC   -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC   OTHER PROTEINS, CAPABLE, IN VITRO, TO UBIQUITINATE HISTONE H2A.
CC   -!- CATALYTIC ACTIVITY: ATP + ubiquitin - protein lysine - AMP +
CC   diphosphate + protein N-ubiquityllysine.
CC   -!- PATHWAY: Ubiquitin conjugation; second step.
CC   -!- M-SCHEANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC   UBIQUITIN-THIOLESTER FORMATION.
CC   -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC   -----
CC   This SWISS-PROI entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: M65083; -; NOT ANNOTATED CDS.
DR   EMBL: U18530; AAB64489.1; ALI_INIT.
DR   FMBL: 228129; CAAB2526.1; -
DR   PIR: A39430; A39430.
DR   HSP: P15731; IQCC.
DR   SGD: S0009736; UBC8.
DR   InterPro: IPR000608; UBCO conjugat.

```

```

DR Pfam: PF00179; UQ_con: 1.
DR ProDom: PD000461; UQ_ConJugat: 1.
DR SMART: SM00212; URCG: 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAL_1: 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAL_2: 1.
KW Ubi conjugation pathway; Ligase; Multigene family.
FI BINDING 85 85 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 218 AA: 24633 MW: 24464884679CE6F CRC64:

Query Match:          9.8%; Score 129; DH 1; Length 218;
Best Local Similarity 22.7%; Pred. No. 0.00047;
Matches 53; Conservative 42; Mismatches 78; Indels 50; Caps 11:

CY 14 VITFEDRMLFSALIKGPTTPYEDGLYFDIQLPNITYPVPPHFCYLSQCSGRLNPNL- 72
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 22 VDLINDSMQGFHVAKLGPKEIPYENCVRHLHVELPDNYPKSPSICGVNKL---FHPNLD 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY 73 YDNGKVCVSL--GTWIGKTERWTSKSSLLQVLISIQGLI--LVNEPYNEAGPDSRGL 129
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 79 IAGSGICLDVNSTW-----SPLYDLINIVEWMIPLGLKEP-----NGSDP---- 119
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY 130 QRGYENSRGVNFMALIRVVQSMQLVRRPPVFEQEYRQHFSTGGWRLVNRJESWCEETHA 189
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 220 -----LNNEATLQL-----RKKLYEKEKE-----YIDKYA 147
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
CY 190 LLEKAQALPNCVPKASSPEPPAVAEASDSGQEPEDG-GPAFGEASQGSDE 241
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 148 TKEKYQOMFGDNDSDSDSGGDLQH-EDSASDEDMIGTCVSSCDSDVDLSE 199
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: April 10, 2003, 10:32:54
 Job time : 17.5862 secs

GenCore version 5.1.4, p5_4578
Copyright (c) 1993 - 2003 Computer Ltd.

QM protein - protein search using sw model

Run on: April 10, 2003, 10:27:50 ; Search time 48.7694 Seconds
(without alignments)
770.497 Million cell: updates/sec

Title: US-09-930-026-2

Perfect score: 1511

Sequence: 1 MAQQQMSKALMLELKL.....MAPOQKPIHSGWNTSSC 282

Scoring table: BLOSUM62

Gapop 10.0 / Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 135 summaries

Database: A:geneseq_101002:*

1: /SID82/gcdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SID82/gcdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SID82/gcdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SID82/gcdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SID82/gcdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SID82/gcdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SID82/gcdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SID82/gcdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SID82/gcdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SID82/gcdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SID82/gcdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SID82/gcdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SID82/gcdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SID82/gcdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SID82/gcdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SID82/gcdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SID82/gcdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SID82/gcdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SID82/gcdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SID82/gcdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SID82/gcdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SID82/gcdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SID82/gcdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1511	100.0	282	AAU97495	Human ubiquitin-co
2	1235	81.7	238	ABH97615	Novel human protei
3	1212	80.2	242	AAH33833	Human secreted pro
4	1212	80.2	232	AAH33834	Human secreted pro
5	1010	66.8	298	AAH99655	Human UBC/CDC34 pr
6	1010	66.8	298	AAH99655	Human UBC/CDC34 pr
7	1010	66.8	298	AAH99655	Human UBC/CDC34 pr
8	1010	66.8	298	AAH99655	Human UBC/CDC34 pr
9	975.5	64.6	193	ABH97614	Novel human protei
10	966.5	64.6	232	AAH38738	Breast and ovarian

11	870	57.6	341	22	ABH34334	Drosophila melanog
12	798	52.8	600	22	ABG07764	Novel human diagno
13	779	51.6	164	23	ABH97616	Novel human protei
14	760.5	50.3	307	22	ABG08168	Novel human diagno
15	755.5	50.0	307	22	ABG07765	Novel human diagno
16	460.5	30.5	170	18	AAH37496	Human skeletal mus
17	460.5	30.5	170	20	AAH27371	Human ubiquitin-co
18	460.5	30.5	170	20	AAH27371	Human HUCE-1 prote
19	460.5	30.5	170	23	AAU97919	Murine ubiquitin b
20	452	29.9	167	21	AAH38559	Arabidopsis thalia
21	452	29.9	167	21	AAH38559	Arabidopsis thalia
22	452	29.9	167	21	AAH38559	Arabidopsis thalia
23	452	29.9	167	21	AAH38559	Arabidopsis thalia
24	450	29.8	166	21	AAH38559	Arabidopsis thalia
25	450	29.8	166	21	AAH38559	Arabidopsis thalia
26	450	29.8	166	21	AAH38559	Arabidopsis thalia
27	450	29.8	166	21	AAH38559	Arabidopsis thalia
28	444	29.4	169	21	AAH38559	Arabidopsis thalia
29	442	29.3	169	21	AAH38559	Arabidopsis thalia
30	439	29.2	166	21	AAH38559	Arabidopsis thalia
31	434.5	28.8	168	22	ABH71776	Drosophila melanog
32	414	27.4	140	21	AAH44743	Arabidopsis thalia
33	414	27.4	140	21	AAH44743	Arabidopsis thalia
34	389	25.7	167	22	ABH60648	Drosophila melanog
35	380.5	25.2	165	20	AAH31983	Human ubiquitin co
36	380.5	25.2	165	20	AAH31983	Human ubiquitin co
37	380.5	25.2	165	20	AAH31983	Human ubiquitin co
38	366	24.2	167	22	AAH70896	Amino acid sequenc
39	362	24.0	125	21	AAH38660	C albicans apoptos
40	362	24.0	125	21	AAH38660	Arabidopsis thalia
41	358.5	23.7	165	22	AAH44744	Arabidopsis thalia
42	334	21.4	98	21	AAH44744	S cerevisiae apopl
43	333	21.4	145	22	AAH01038	Arabidopsis thalia
44	309	20.5	109	21	AAH09144	Human polypeptide
45	309	20.5	109	21	AAH09144	Arabidopsis thalia
46	301	19.9	109	21	AAH41296	Arabidopsis thalia
47	299	19.8	109	21	AAH41296	Arabidopsis thalia
48	297	19.5	152	16	AAH43423	Zea mays protein f
49	295	19.5	152	16	AAH79654	Human cancer assoc
50	295	19.5	152	20	AAH39967	Human E2 protein..
51	295	19.5	152	21	AAH03175	Human rad6 homolog
52	295	19.5	152	21	AAH03175	Human rad6 homolog
53	295	19.5	152	21	AAH03175	Human ovarian anti
54	294	19.5	152	21	AAH03175	Human colon cancer
55	294	19.5	152	21	AAH03175	Zea mays protein f
56	280	18.5	152	21	AAH35424	Zea mays protein f
57	277	18.3	151	22	ABH58757	Arabidopsis thalia
58	276	18.3	151	22	ABH58757	Drosophila melanog
59	276	18.3	151	22	ABH58757	Ubiquitin-conjugat
60	272	18.0	152	21	AAH0240	Arabidopsis thalia
61	272	18.0	152	21	AAH0240	Arabidopsis thalia
62	271	17.9	152	21	AAH0240	Zea mays Rad6 prot
63	270	17.9	152	21	AAH0240	Zea mays protein f
64	268	17.7	172	23	ABH42174	Zea mays Rad6 prot
65	266	17.6	143	21	AAH32531	Human ovarian anti
66	266	17.6	143	21	AAH32531	Arabidopsis thalia
67	266	17.6	150	21	AAH15020	Arabidopsis thalia
68	266	17.6	150	21	AAH47818	Arabidopsis thalia
69	264.5	17.5	177	18	AAH31278	Arabidopsis thalia
70	262	17.3	143	21	AAH10241	Clam cyclin-selec
71	261	17.3	143	21	AAH37254	Arabidopsis thalia
72	259	17.1	148	21	AAH70092	Arabidopsis thalia
73	259	17.1	148	21	AAH70092	Zea mays Rad6 prot
74	257	17.0	143	21	AAH35168	Zea mays protein f
75	254.5	16.8	173	23	ABH41313	Human ovarian anti
76	254.5	16.8	177	18	AAH31280	Clam cyclin-selec
77	254.5	16.8	179	20	AAH31277	Human cyclin-selec
78	254.5	16.8	179	20	AAH75655	A cyclin related p
79	254.5	16.8	192	23	AAH43461	Protein UCHL10 dif
80	254	16.8	172	21	AAH26041	Zea mays protein f
81	252	16.7	141	21	AAH15021	Arabidopsis thalia
82	252	16.7	141	21	AAH15021	Arabidopsis thalia
83	252	16.7	148	21	AAH32309	Arabidopsis thalia

84 251.5 16.6 148 21 AAG59017 Arabidopsis thalia
85 251.5 16.6 181 23 AB937609 Novel human protei
86 251 148 21 AAG13466 Arabidopsis thalia
87 251 148 21 AAG26042 Zea mays protein f
88 251 148 21 AAG37711 Arabidopsis thalia
89 250 148 21 AAG47509 Arabidopsis thalia
90 249 147 21 AAG34325 Zea mays protein f
91 248.5 16.4 129 21 AAG32532 Arabidopsis thalia
92 248.5 16.4 156 21 AAG27041 Zea mays protein f
93 248 147 21 AAG34344 Zea mays protein f
94 247 148 21 AAY70094 Zea mays Rad6 prot
95 245 148 21 AAG3237 Zea mays protein f
96 245 16.2 148 17 AAR39264 Cotton plant ubiq
97 244.5 16.2 119 21 AAG27042 Zea mays protein f
98 244.5 16.2 179 18 AAW11279 Human cyclin-selec
99 243.5 16.1 119 21 AAG35169 Zea mays protein f
100 240.5 15.9 151 22 ABB65427 Drosophila melanog
101 240.5 15.9 151 22 AAB51234 Ubiquitin-conjugat
102 240 90 21 AAG32681 Zea mays protein f
103 240 15.9 90 21 AAG35425 Zea mays protein f
104 240 15.9 112 21 AAGA1092 Zea mays protein f
105 239 15.8 178 22 ABB72000 Drosophila melanog
106 239 15.8 178 23 AAM47580 Drosophila cell cy
107 238 15.8 110 21 AAG26116 Zea mays protein f
108 237 15.7 147 22 ABB63246 Drosophila melanog
109 237 15.7 147 22 ABB63256 Drosophila melanog
110 237 15.7 148 21 AAG08607 Arabidopsis thalia
111 237 15.7 148 21 AAG43234 Arabidopsis thalia
112 235.5 15.6 147 16 AAR79645 C. albicans caubCE
113 235.5 15.6 147 20 AAY39962 Candida albicans
114 235.5 15.6 147 21 AAB03170 Human TIABP1. Hom
115 235.5 15.6 158 16 AAR22829 Ubiquitin-conjugati
116 235.5 15.6 158 17 AAW04240 Rat UBCE2A. Rattu
117 235.5 15.6 158 18 AAM12382 Human ubiquitin co
118 235.5 15.6 158 18 AAM14135 Human ubiquitin-co
119 235.5 15.6 158 19 AAW64716 Human TIABP1 prote
120 235.5 15.6 158 20 AAY39964 Human ubiquitin co
121 235.5 15.6 158 21 AAB19281 A frog ubiquitin-c
122 235.5 15.6 158 21 AAB03173 Human ubiquitin-co
123 235.5 15.6 158 23 ABB79002 Human UBE2I protei
124 235.5 15.6 205 21 AAB58187 Lung cancer associ
125 235.5 15.6 158 16 AAM11944 p53 binding protei
126 234.5 15.5 158 18 AAR79652 Human rapUBC enzy
127 234.5 15.5 158 19 AAW64718 Human TIABP1 genom
128 234.5 15.5 158 19 AAW57333 Human WBP1 protein
129 234.5 15.5 354 22 ABB65272 Drosophila melanog
130 234.5 15.5 117 21 AAG15622 Arabidopsis thalia
131 234.5 15.5 160 23 ABB42265 Human ovarian anti
132 233.5 15.5 147 22 AAR92968 Human protein sequ
133 233 15.4 155 23 ABB77472 Fungi stress respo

ALIGNMENTS

RESULT 1
AAU97495
XX AAU97495 standard; Protein: 282 AA.
AC AAU97495;
XX AAU97495;
XX 13-AUG-2002 (first entry)
XX Human ubiquitin-conjugating enzyme homologue UCEH2.
XX Human; antiinflammatory; immunomodulator;
XX autolysosomal disease; cytoskeletal; neurotropic;
XX immunosuppressive.
XX Homo sapiens.
XX

PN US6277568-B1.
XX 21-AUG-2001.
XX 09-APR-1998; 98US-0058368.
XX 09-APR-1998; 98US-0058368.
XX (INCY-) INCYTE GENOMICS INC.
XX Lal P, Hallman JL, Guegler KJ, Corley NC, Raughn M, Azimzai Y;
XX WPI: 2002-433227/46.
XX N-PSDB; ABB52043.
XX New polynucleotide encoding human ubiquitin-conjugating enzyme, useful
XX e.g. for treatment, prevention and diagnosis of cancer, or autoimmune
XX and neuronal disease.
XX Claim 1: Column 41-44; 26pp; English.
XX The present invention relates to a new polynucleotide that encodes a
XX polypeptide with ubiquitin-conjugating activity. The polynucleotide of
XX the invention encodes the human ubiquitin-conjugating enzyme UCEH2 and
XX is used for recombinant expression of this enzyme, as primers and probes,
XX e.g. for diagnosis, monitoring or gene mapping and as a source of
XX ribozymes, antisense sequences or triplex-forming therapeutic agents.
XX Polypeptides encoded by the invention, and optionally expressed from
XX gene therapy vectors, are useful for treatment, diagnosis and prevention
XX of a wide variety of cancers and autoimmune and neuronal diseases.
XX The present amino acid sequence represents the human
XX ubiquitin-conjugating enzyme homologue UCEH2 protein of the invention.
XX Sequence 282 AA;
XX
XX Query Match 100.0%; Score 1511; DB 23; Length 282;
XX Best Local Similarity 100.0%; Pred. No. 2.8e-147;
XX Matches 282; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 1 MAQQQMTSSQKALMELKSLQEPVEGRITLVDESDLYNNEVAIEGLPTLYEGGYFKA 60
Db 1 MAQQQMTSSQKALMELKSLQEPVEGRITLVDESDLYNNEVAIEGLPTLYEGGYFKA 60
QY 61 HIKEPIDYSPPTFRFLTKMHPNIEYNGDWICISILIPVDPQSGELPSERNWPTQNV 120
Db 61 HIKEPIDYSPPTFRFLTKMHPNIEYNGDWICISILIPVDPQSGELPSERNWPTQNV 120
QY 121 RTILSVLSLLNEPNTFSPANDASVMEKWRDSKGDKEYAEIIRKQVSATKAPEKDG 180
Db 121 RTILSVLSLLNEPNTFSPANDASVMEKWRDSKGDKEYAEIIRKQVSATKAPEKDG 180
QY 181 VKVPTTILAEYCIKTKVPSNDSDLLYDDLYDDDDIDDDHDFEDALCYDDDDSCMRSDV 240
Db 181 VKVPTTILAEYCIKTKVPSNDSDLLYDDLYDDDDIDDDHDFEDALCYDDDDSCMRSDV 240
QY 241 LQCPCTALPSQAKRGASGDLAMAPQOKPIHSGWGNTHSSC 282
Db 241 LQCPCTALPSQAKRGASGDLAMAPQOKPIHSGWGNTHSSC 282
RESULT 2
ABB97615
XX ABB97615 standard; Protein: 238 AA.
XX ABB97615;
XX ABB97615;
XX 27-JUN-2002 (first entry)
XX Novel human protein SEQ ID NO: 883.
XX Human; antianemic; vulnery; antiinflammatory; immunomodulator;
XX antiinfective; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX

expressed sequence tag.
 XX Homo sapiens.
 OS W0200222660-A2.
 PN W0200222660-A2.
 PD 21-MAR-2002.
 XX 10-SEP-2000; 2000US-0659671.
 XX 11-SEP-2000; 2000US-0659671.
 PR (HYSE-) HYSEQ INC.
 PA Tang Y., Liu C., Zhou P., Asundi V., Zhang J., Zhao Q., Ren F.,
 PI Xue A.J., Yang Y., Wehrman I., Ormanac R.;
 XX WPI: 2002-292408/33.
 DR N-PSDB; ABN32802.
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 PS Example 2: SEQ ID NO 883; 509pp; English.
 XX The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate active or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention.
 XX Sequence 238 AA:
 SQ
 Query Match 91.7%; Score 1235; DB 23; Length 238;
 Best Local Similarity 99.1%; Pred. No. 6.7e-119;
 Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAQQQMTSSKALMLELKSLEQEPVEGFRITLVDESOLYNWEVAIFGLPNTLYEGGYFKA 60
 DB 1 MAQQQMTSSKALMLELKSLEQEPVEGFRITLVDESOLYNWEVAIFGLPNTLYEGGYFKA 60
 QY 61 MKKPPIDYPSPTFRFLTKMHPNIYENGVCISILHPVDDPQSGELPSERNPTQNV 120
 DB 61 MKKPPIDYPSPTFRFLTKMHPNIYENGVCISILHPVDDPQSGELPSERNPTQNV 120
 QY 121 RTILLSVISLNEPNTFSPANVDASVHFKKWRDSKCKKKEYAEIIRKQVSATKAEAKDG 180
 DB 121 RTILLSVISLNEPNTFSPANVDASVHFKKWRDSKCKKKEYAEIIRKQVSATKAEAKDG 180
 QY 181 VKVFTIARCIKTKVPSNDSSDLLYDDLYDDDDDEEEDADCYDDDDSG 234
 DB 181 VKVFTIARCIKTKVPSNDSSDLLYDDLYDDDDDEEEDADCYDDDDSG 234
 RESULT 3
 AAB33833
 ID AAB33833 standard; Protein: 232 AA.
 XX
 AC AAB33833;
 XX
 DT 02-FEB-2001 (first entry)
 XX Human secreted protein BLAST search protein SEQ ID NO: 177.
 DE
 DE Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiact; gene therapy; cancer; immune disorder; cardiovascular disorder;

neurological disease; infection; human; secreted protein.
 XX Homo sapiens.
 OS W0200056753-A1.
 PN W0200056753-A1.
 PD 28-SEP-2000.
 XX 16-MAR-2000; 2000WO-US06765.
 XX 23-MAR-1999; 99US-0126051.
 PR 10-DEC-1999; 99US-0169906.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM, Komatsoulis G;
 PI WPI: 2000-594570/56.
 DR Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder -
 PS Disclosure: Page 423-424; 410pp; English.
 XX The invention relates to the isolation of genes AAC59277-C59325 encoding
 CC the human secreted proteins AAB33718-B33764. The sequence is a search
 CC result from a BLASTX homology search. The genes and proteins are useful
 CC for preventing, ameliorating or treating medical conditions, e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer, and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 XX Sequence 232 AA:
 SQ
 Query Match 80.2%; Score 1212; DB 21; Length 232;
 Best Local Similarity 99.6%; Pred. No. 1.5e-116;
 Matches 227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 7 TSSOKALMLELKSLEQEPVEGFRITLVDESOLYNWEVAIFGLPNTLYEGGYFKAHKFPPI 66
 DB 1 TSSOKALMLELKSLEQEPVEGFRITLVDESOLYNWEVAIFGLPNTLYEGGYFKAHKFPPI 66
 QY 67 DYPYSPPTFRFLTKMHPNIYENGVCISILHPVDDPQSGELPSERNPTQNVRIILLS 126
 DB 67 DYPYSPPTFRFLTKMHPNIYENGVCISILHPVDDPQSGELPSERNPTQNVRIILLS 126
 QY 127 VISLILNEPNTFSPANVDASVHFKKWRDSKCKKKEYAEIIRKQVSATKAEAKDGKVPVPT 186
 DB 127 VISLILNEPNTFSPANVDASVHFKKWRDSKCKKKEYAEIIRKQVSATKAEAKDGKVPVPT 186
 QY 187 LAEYCIKTKVPSNDSSDLLYDDLYDDDDDEEEDADCYDDDDSG 234
 DB 187 LAEYCIKTKVPSNDSSDLLYDDLYDDDDDEEEDADCYDDDDSG 234
 RESULT 4
 AAB33834
 ID AAB33834 standard; Protein: 232 AA.
 XX
 AC AAB33834;
 XX
 DT 02-FEB-2001 (first entry)
 XX

DE Human secreted protein: BLAST search: protein SEQ ID NO: 178.
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vimentary; actinconvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX Homo sapiens.
 OS
 XX WC2000056753-A1.
 XX 28-SEP-2000.
 PD
 XX 16-MAR-2000; 2000WC-US06765.
 PF
 XX 23-MAR-1999; 94US-012605.
 PA 10-DEC-1999; 94US-0169506.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM, Komatsolis G;
 PI WPI: 2000-594570/56.
 XX Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder -
 XX Disclosure: Page 424-425; 40pp; English.
 PS
 XX The invention relates to the isolation of genes AAC59277-C59325 encoding
 CC the human secreted proteins AAB33718-B33764. The sequence is used as a
 CC query sequence for doing BLASTX searches to identify homologous
 CC sequences. The genes and proteins are useful for preventing,
 CC ameliorating or treating medical conditions, e.g. by protein or gene
 CC therapy. The genes are isolated from a range of human tissues disclosed
 CC in the specification, the nucleic acids, proteins, antibodies and
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal
 CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
 CC disease, allergies, autoimmune haemolytic anaemia, autoimmune
 CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
 CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular
 CC disorders such as myocardial ischaemias; (d) wound healing; (e)
 CC neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections.
 XX Sequence 232 AA:
 QY 7 TSSQKALMELKSLQEPVEFRITLVDSDLYNNEVAIFGLPNTLYEGGYFAHKIFPI 66
 DB 1 TSSQKALMELKSLQEPVEFRITLVDSDLYNNEVAIFGLPNTLYEGGYFAHKIFPI 60
 QY 67 DYPSPTPTFRFLTKMHPNIVENSVCISILHPPVDDPQSGELPSERNWNTQNY 126
 DB 61 DYPSPTPTFRFLTKMHPNIVENSVCISILHPPVDDPQSGELPSERNWNTQNY 120
 QY 127 VISLNLNPNFSPANVDASVMFKWRKSGKKEYAEIRKQVSAIKAEAKDGVKVPIT 186
 DB 121 VISLNLNPNFSPANVDASVMFKWRKSGKKEYAEIRKQVSAIKAEAKDGVKVPIT 180
 QY 187 LAEYCIKTKVPSNDNSLLYDGLYDDIDDEDFEEDADCYDDDDSG 234
 DB 181 LAEYCIKTKVPSNDNSLLYDGLYDDIDDEDFEEDADCYDDDDSG 228
 RESULT 5

AAR79655
 ID AAR79655 standard; Protein; 298 AA.
 XX
 AC AAR79655;
 XX
 DT 06-DEC-1995 (first entry)
 DE
 XX Human UBC/CDC34 protein.
 XX
 KW Ubiquitin-conjugating enzyme; UBC/CDC34; cell cycle;
 KW cell proliferation; cancer; psoriasis; fibrosis.
 XX
 OS Homo sapiens.
 XX W09518974-A.
 PN 13-JUL-1995.
 XX
 XX 04-JAN-1995; 95WC-US00164.
 PF
 XX 13-SEP-1994; 94US-0305520.
 PR 04-JAN-1994; 94US-0176937.
 PR 23-MAY-1994; 94US-0247904.
 PR 27-MAY-1994; 94US-0250795.
 XX
 PA (MITO-) MITOTIX INC.
 XX
 XX Cottarel G, Draetta G, Eckstein JW, Gyuris J, Roife M;
 PI WPI: 1995-255137/33.
 DR N-PSDB; AAR97845.
 XX
 PT Identifying inhibitors of ubiquitin mediated proteolysis of cell cycle
 PT regulatory proteins - also new ubiquitin conjugating enzymes, their
 PI related nucleic acid, vectors, antibodies etc., useful for regulating
 PI e.g. cell proliferation
 XX
 PS Disclosure: Page 107-108; 157pp; English.
 XX
 CC Human UBC3/CDC34 cDNA (given in AAR97845) was amplified from a HeLa
 CC cell cDNA library. The gene was subcloned into a baculovirus or
 CC pGEX vector for expression of recombinant UBC/CDC34 in Sf9 insect
 CC or E. coli cells for use as a component of an in vitro ubiquitin
 CC conjugating system.
 XX
 SQ Sequence 298 AA:
 QY 1 MAQOQMTSSQKALMELKSLQEPVEFRITLVDSDLYNNEVAIFGLPNTLYEGGYFA 60
 DB 63 MARPLVPSQKALLLELKLQEPVEFRITLVDSDLYNNEVAIFGLPNTLYEGGYFA 122
 QY 61 HIKFPIDYPSPTFRFLTKMHPNIVENSVCISILHPPVDDPQSGELPSERNWNTQNY 120
 DB 123 RLKFPIDYPSPTFRFLTKMHPNIVENSVCISILHPPVDDPQSGELPSERNWNTQNY 182
 QY 121 RTLLSVISLNLNPNFSPANVDASVMFKWRKSGKKEYAEIRKQVSAIKAEAKD 180
 DB 183 RTLLSVISLNLNPNFSPANVDASVMFKWRKSGKKEYAEIRKQVSAIKAEAKD 242
 QY 181 VKVPTTLAEYCIKTKVPSNDNSLLYDGLYDDIDDEDFEEDAD-CY--DDDDSG 234
 DB 243 VKVPTTLAEYCIKTKVPSNDNSLLYDGLYDDIDDEDFEEDAD-CY--DDDDSG 294
 RESULT 6
 AAY39971
 ID AAY39971 standard; Protein; 298 AA.
 XX
 AC AAY39971;

CC epidermal conditions such as psoriasis, neoplastic epidermal conditions,
 CC skin cancers e.g., basal cell carcinomas, squamous cell carcinomas. The
 CC inhibitors are useful for regulating myc expression and rendering the
 CC cells sensitive to chemotherapeutic treatment or to upset the balance of
 CC transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-
 CC mediated degradation of cyclins are useful as antiproliferative agents.
 CC The present sequence represents human: UBC/CDC34.
 XX
 SO Sequence 298 AA;

Query Match 66.8%; Score 1010; DB 21; Length 298;
 Best Local Similarity 79.7%; Pred. No. 1.6e-95;
 Matches 189; Conservative 18; Mismatches 22; Indels 8; Gaps 3;

OY 1 MAQOMTSQKALMELKSGEPEVGRITLVDSGLYNNEVAIFGLPNTLYEGGYEKA 60
 DB 63 MARPLVPSQKALLLEKGLQEPVGEFVILVEGOLYNNEVAIFGPPNTYEGGYEKA 122
 OY 61 HIKFPIDYPSPTFERELIKMHPNIYENGVCISILHPVDDPQSGELPSRWNPNTNV 120
 DB 123 ELKFPIDYPSPTFERELIKMHPNIYENGVCISILHPVDDPQSGELPSRWNPNTNV 182
 OY 121 RTILSVISLLNEPNTSPANVDASVMFKKRSKDKGKVEAEIIRKQVSAKAEKDG 180
 DB 183 RTILSVISLLNEPNTSPANVDASVMFKKRSKDKGKVEAEIIRKQVSAKAEKDG 242
 OY 181 VKVPTTILAEYCIKTKVPSNDSSLLYDDLYDDEDEDEEDAD-CY--DDDSG 234
 DB 243 VKVPTTILAEYCVK-KAPAPDEGSLFYDDYED-----GEVEEADSCFGDEDDSG 294

RESULT 8
 AA084320
 ID: AA084320 standard; Protein: 298 AA.
 XX
 AC AA084320;

DT 08-MAY-2002 (first entry)

XX Protein CDC34 differentially expressed in breast cancer tissue.

XX Human: diagnosis of breast cancer; endometrial cancer; breast tumour;
 KW MAI: mitotic activity index; cytostatic.
 XX

OS Homo sapiens.

XX W200210436-A2.

XX W200210436-A2.

XX 07-FEB-2002.

XX 27-JUL-2001; 2001WO-US23642.

XX 23-JUL-2000; 2000US-222693P.

XX (BOHM) BRIGHAM & WOMENS HOSPITAL INC.

XX (BAAK/) BAAK J.

XX Baak J, Mitter GL.

XX WPI: 2002-180084/23.

XX N-PSDB: ABK35540.

XX Diagnosing breast cancer comprises determining expression of nucleic
 PT acid molecules or expression products that are differentially expressed
 PT in normal and malignant tissue -
 XX
 PS Claim 37; Page 131-132; 219pp; English.

XX The present invention relates to a method for diagnosing breast
 CC in a subject suspected of having endometrial cancer. The method
 CC comprises determining the expression of a set of human genes or
 CC expression products in an endometrial sample suspected of being
 CC cancerous. The human genes of the invention are differentially
 CC

CC expressed in breast tumours characterised as high or low MAI (mitotic
 CC activity index). These sets of genes can be used to discriminate between
 CC high and low MAI breast tumours. The invention also provides DNA and
 CC protein microarrays for analysing the expression of the human genes and
 CC their protein products. The methods and arrays are useful for the
 CC diagnosis and prognosis of endometrial cancer, selecting and monitoring
 CC treatment regimes, and identification of compounds useful for the
 CC treatment of endometrial cancer. AA084311-AA084361 represent the human
 CC proteins of the invention that are differentially expressed in breast
 CC cancer tissue.
 XX

SO Sequence 298 AA;

Query Match 66.8%; Score 1010; DB 23; Length 298;
 Best Local Similarity 79.7%; Pred. No. 1.6e-95;
 Matches 189; Conservative 18; Mismatches 22; Indels 8; Gaps 3;

OY 1 MAQOMTSQKALMELKSGEPEVGRITLVDSGLYNNEVAIFGLPNTLYEGGYEKA 60
 DB 63 MARPLVPSQKALLLEKGLQEPVGEFVILVEGOLYNNEVAIFGPPNTYEGGYEKA 122
 OY 61 HIKFPIDYPSPTFERELIKMHPNIYENGVCISILHPVDDPQSGELPSRWNPNTNV 120
 DB 123 ELKFPIDYPSPTFERELIKMHPNIYENGVCISILHPVDDPQSGELPSRWNPNTNV 182
 OY 121 RTILSVISLLNEPNTSPANVDASVMFKKRSKDKGKVEAEIIRKQVSAKAEKDG 180
 DB 183 RTILSVISLLNEPNTSPANVDASVMFKKRSKDKGKVEAEIIRKQVSAKAEKDG 242
 OY 181 VKVPTTILAEYCIKTKVPSNDSSLLYDDLYDDEDEDEEDAD-CY--DDDSG 234
 DB 243 VKVPTTILAEYCVK-KAPAPDEGSLFYDDYED-----GEVEEADSCFGDEDDSG 294

RESULT 9
 AA084320
 ID: AA084320 standard; Protein: 193 AA.
 XX
 AC AA084320;

DT 27-JUN-2002 (first entry)

XX Novel human protein SEQ ID NO: 882.

XX Human: antitumour; antineoplastic; immunomodulator;
 KW antineoplastic; cytostatic; cytotoxic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; ESI;
 XX expressed sequence tag.

XX Homo sapiens.

XX W0200222660-A2.

XX 21-MAR-2002.

XX 10-SEP-2001; 2001WO-US26015.

XX 11-SEP-2000; 2000US-0659671.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI: 2002-292408/33.

XX N-PSDB: ABN32800.

XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 XX
 PS Example 2; SEQ ID NO 882; 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC hematopoiesis e.g. to treat aplastic anemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention.
 XX

XX Sequence 193 AA:

Query Match 64.6%, Score 975.5; DB 23; Length 193;
 Best Local Similarity 79.9%; Pred. No. 3e-92;
 Matches 187; Conservative 0; Mismatches 2; Indels 45; Gaps 1;

QY 1 MAQQMTSSQKALMLKLSIQEPEVEGFRITLVDSIDLYNNEVAIFGIPNTLYEGGYFA 60
 DB 1 MAQQMTSSQKALMLKLSIQEPEVEGFRITLVDSIDLYNNEVAIFGIPNTLYEGGYFA 60
 QY 61 HIREPIDYPSPTTREFCTKMHHPNIYENGDCVISTILHPVDDPQSGELPSERNWPTQNV 120
 DB 61 HIREPIDYPSPTTREFCTKMHHPNIYENGDCVISTILHPVDDPQSGELPSERNWPTQNV 120
 QY 121 RTLLSVISLNEPNTFSPANVDASVWFKKWSKDKKEYAETIRKOVSAKAEKDG 180
 DB 121 RTLLSVISLNEPNTFSPANVDASVWFKKWSKDKKEYAETIRKOVSAKAEKDG 180

QY 181 VKYPTTLAEYCIKTKVPSNNSLLYDDCYDDDDDEDEEREDADCYDDDDSG 234

DB 136 VKYPTTLAEYCIKTKVPSNNSLLYDDCYDDDDDEDEEREDADCYDDDDSG 234

RESULT 10

AA858748

ID AA858738 standard; Protein: 232 AA.

XX AA858738:

XX 27-MAR-2001 (first entry)

DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 446.

XX Human: breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW nototropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antitumor; antitumor; antitumor;
 KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.

XX Homo sapiens.

XX W0200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 200000 US05881.

XX 12-MAR-1999; 4905-012427C.

XX (HOMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM.

XX W01: 2000-011515/58.

XX N PSOR; AAF21641.

XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases.

XX Claim 11; Page 874-875; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAF58711 - AAF59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAF59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC nototropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antitumor; antitumor; antitumor;
 CC antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemia; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 232 AA:

Query Match 64.0%, Score 966.5; DB 21; Length 232;
 Best Local Similarity 79.6%; Pred. No. 3.4e-91;
 Matches 183; Conservative 16; Mismatches 22; Indels 9; Gaps 4;

QY 8 SSOKALMLKLSIQEPEVEGFRITLVDSIDLYNNEVAIFGIPNTLYEGGYFAKHKPTD 67
 DB 5 SSOKALMLKLSIQEPEVEGFRITLVDSIDLYNNEVAIFGIPNTLYEGGYFAKHKPTD 64
 QY 68 YPSPPTREFCTKMHHPNIYENGDCVISTILHPVDDPQSGELPSERNWPTQNV 127
 DB 65 YPSPPTREFCTKMHHPNIYENGDCVISTILHPVDDPQSGELPSERNWPTQNV 124
 QY 128 TSLNEPNTFSPANVDASVWFKKWSKDKKEYAETIRKOVSAKAEKDGKVKVPTTL 187
 DB 125 TSLNEPNTFSPANVDASVWFKKWSKDKKEYAETIRKOVSAKAEKDGKVKVPTTL 183

XX Sequence 232 AA:

Query Match 64.0%, Score 966.5; DB 21; Length 232;
 Best Local Similarity 79.6%; Pred. No. 3.4e-91;
 Matches 183; Conservative 16; Mismatches 22; Indels 9; Gaps 4;

QY 8 SSOKALMLKLSIQEPEVEGFRITLVDSIDLYNNEVAIFGIPNTLYEGGYFAKHKPTD 67

DB 5 SSOKALMLKLSIQEPEVEGFRITLVDSIDLYNNEVAIFGIPNTLYEGGYFAKHKPTD 64

QY 68 YPSPPTREFCTKMHHPNIYENGDCVISTILHPVDDPQSGELPSERNWPTQNV 127

DB 65 YPSPPTREFCTKMHHPNIYENGDCVISTILHPVDDPQSGELPSERNWPTQNV 124

QY 128 TSLNEPNTFSPANVDASVWFKKWSKDKKEYAETIRKOVSAKAEKDGKVKVPTTL 187

DB 125 TSLNEPNTFSPANVDASVWFKKWSKDKKEYAETIRKOVSAKAEKDGKVKVPTTL 183

QY 188 AYCITKTPSNNSLLYDDCYDDDDDEDEEREDADCYDDDDSG 234

DB 184 AYCITKTPSNNSLLYDDCYDDDDDEDEEREDADCYDDDDSG 228

XX RESULT 11

XX ABB63434

XX ID ABB63434 standard; Protein: 341 AA.

XX AC ABB63434;

XX DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

XX Drosophila melanogaster polypeptide SEQ ID NO 17094.

QY 16' -----KQVSATKAE 176
 DB 48: AAASFLSVKQCHAAHMDCAIPGSPASSNGCTFTRVLEVPKAGSSKQVLGKQDA 540
 QY 177 EKDKVPTTLAEYCTKTKVPSNDSLLYDLYDQIDDEDEEDAD-CY--DDGDS 233
 DB 541 ERDGVKPTTLAEYCVKTKAPGEGSLFYDYED-----GEVEEACSCFDDDDDS 595
 QY 234 G 234
 DB 596 G 596

RESULT 13
 ABH97616
 ID ABH97616 standard; Protein: 164 AA.
 AC ABH97616;
 DT 27-JUN-2002 (first entry)
 DE Novel human protein SEQ ID NO: 884.
 KW Human; anti-neoplastic; vulcercary; anti-inflammatory; immunomodulator;
 KW anti-fertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; anti-parkinsonian; protein therapy; EST;
 KW expressed sequence tag.
 CS Homo sapiens.
 LN WO200222650-A2.
 PD 21-MAR-2002.
 PR 10-SEP-2001; 2001WO-US26015.
 PR 11-SEP-2000; 2000US-0659671.
 PA (HYSE-) HYSEQ INC.
 PI Tang YI, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AI, Yang Y, Wehrman T, Drmanac RT;
 DR WPI: 2002-292408/33.
 DR N-PSDB; ARN32802.
 PT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 PS Example 2; SEQ ID NO 884; 509pp; English.

The present invention provides the protein and coding sequences of 444 novel human proteins. These were isolated from expressed sequences tags (ESTs). They can be used to stimulate cell growth, to regulate hematopoiesis e.g. to treat aplastic anemia, to help tissue regrowth e.g. in burn treatment, to regulate the immune system e.g. to treat multiple sclerosis, to regulate activin or inhibin e.g. to treat infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke and cancer, to screen for drugs, to treat inflammatory conditions e.g. rheumatoid arthritis, and to treat nervous system disorders e.g. Parkinson's disease. The present sequence is a protein of the invention.

Query Match 51.6%; Score 779; DB 23; Length 164;
 Best Local Similarity 67.5%; Pred. No. 4.5e-72;
 Matches 156; Conservative C; Mismatches 2; Index 74; Gaps 2;

QY 1 MAQQQTSSQKALMLEKSLQEEFVRCFRTLVQESDLYNNEVALFGLPNTHYEGGYFKA 60
 DB 1 MAQQQTSSQKALMLEKSLQEEFVRCFRTLVQESDLYNNEVALFGLPNTHYEGGYFKA 59

QY 61 HIKFPDYFYSPTFRFLAKMMHPNLYENGQVCISILHPPVDDPQSGEPLSRNPTQNV 120
 DB 60 -----NGDVCISILHPPVDDPQSGELPSRNPNTQNV 91
 QY 121 RTILLSVISLLNEINIFSPANDASVMPFRKWRDSKGKDKYAEIHKQVSATKAEKDG 180
 DB 92 -----RKQVSATKAEKDG 106
 QY 181 VKVPTTLAEYCIKTKVPSNDSLLYDLYDQIDDEDEEDADCYDDDSG 234
 DB 107 VKVPTTLAEYCIKTKVPSNDSLLYDLYDQIDDEDEEDADCYDDDSG 160

RESULT 14
 ABC08168
 ID ABC08168 standard; Protein: 307 AA.
 AC ABC08168;
 DT 13-FEB-2002 (first entry)
 DE Novel human diagnostic protein #8159.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 PN WO200075067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YI;
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS72355.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID No 38527; 103pp; English.

The invention relates to isolated polynucleotide (i) and polypeptide (ii) sequences. (i) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (ii). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (i) is useful in gene therapy techniques to restore normal activity of (ii) or to treat disease states involving (ii). (ii) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (ii) and its binding partners are useful in medical imaging of sites expressing (ii). (i) and (ii) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ARG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.


```

CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AHC00010-AHC30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 307 AA;

Query Match 50.0%; Score 755.5; DB 22; Length 307;
Best Local Similarity 64.9%; Pred. No. 3.2e-69;
Matches 161; Conservative 19; Mismatches 47; Indels 21; Gaps 11;

QY : MAQQOMTSQKALMGLKSLQPEPVEGFRPIIVDESDLYNHWVAIFGLPNIIYEGGYFKA 60
   || : || || || || || || || || || || || || || || || || || || || ||
Db 63 MARPLVPSSHKALLLELKGIOEPVEGFRVTVDEGDLYNHWVAIFGPPNIYEGGYFKA 122

QY 61 HKFPIDYSPPTFRFLTKMHPHNIYENGDVCIISLHPFVD-DPQSGELPSERNWPTQN 129
   : || || : || || || || || || || || || || || || || || || || || || ||
Db 123 RUKFPIDYSPAPFAFLTKMHPHNIYETGVCTISLHPFVDTDQSGELPSERNWPTQN 182

QY 120 VRTILLSVTSLNE--PNIF-SPANVDAS--VNRKW-HDSKGDKE-YAELLTKOVSAT 172
   || || || || || || || || || || || || || || || || || || || || ||
Db 183 VDHSSXYXSSLLERRPNPIJRPANVGRLPMLQGSWEREQGGRIRLHRTSFRKQVLGT 242

QY 173 KAEAKGCVKVPITILAFYCIK---TKVPSNDSNLSLLYDDLVDGDFEERDAD-CY 228
   || : || || || || || || || || || || || || || || || || || || || ||
Db 243 KYDAERDGVKVPITILPEYCVKQGGRAPI--DEGSDIFYDDYED-----AAVEERADSCF 295

QY 229 --DDDDSG 234
   || || || ||
Db 296 GDDEDDSG 303

RESULT 16
AW37496
ID AW37496 standard; Protein; 170 AA.
AC
AC AW37496;
AD
DT 20-APR-1998 (first entry)
DE
DE Human skeletal muscle specific ubiquitin conjugating enzyme.
XX
XX Human: foetal brain cDNA library; GDP dissociation stimulating protein;
KW brain specific nucleosome assembly protein; diagnosis; therapy;
KW skeletal muscle specific ubiquitin conjugating enzyme; IMP-2; NPIK;
KW nel-related protein type 1; nel-related type 2; hereditary disease;
XX cancer.
XX
XX Homo sapiens.
OS
OS
PN EP796913-A2.
XX
XX
XX 24-SEP-1997.
PD
PD
XX 19-MAR-1997; 97EP-0104842.
PX
PX
PR 05-MAR-1997; 97JP-0069163.
PR 19-MAR-1996; 96JP-0363410.
XX
XX (SAKA ) OTSUKA PHARM CO LTD.
XX
XX Fujiwara T, Horie M, Watanabe T;
XX WPI: 1997-459830/43

```

DR N-PSDB: AAV01572, AAV01873.
 XX Novel human genes, e.g. brain-specific nucleosome assembly protein -
 PT useful for diagnosis or therapy of hereditary disease and cancer
 XX
 PS Claim 7: Page 64-65; 123pp; English.
 XX
 CC The present sequence represents a skeletal muscle specific ubiquitin
 CC conjugating enzyme isolated from a human foetal brain cDNA library. The
 CC nucleotide or amino acid sequences are useful for in-vitro diagnosis of
 CC hereditary diseases and cancer and for preparation of pharmaceuticals.
 XX
 SQ Sequence 170 AA:
 Query Match 30.58; Score 460.5; DB 18; Length 170;
 Best Local Similarity 52.48; Pred. No. 3.7e-39;
 Matches 87; Conservative 28; Mismatches 44; Indels 7; Gaps 3;
 QY 6 MISSOKALM:--ELKSLOEPPVEGFRITLVDSDLYNNEVAIFGLPNTLYEGGYFKAHIK 63
 DB 1 MTEIQSALLRLRQIAELKNKPVGFSAGLIDDDLYNNEV:ITGPPDTLYEGGVFKAHLT 60
 QY 64 FPIDYPSPTFRFLTKMWHENIYENGDCVCSILHPPVDDPQSGELPSEWNPQWRTI 123
 DB 61 FKDYPLRPKMKFTETWHPNVDKNSDVCISILHEPGEDKYGYEKPEERWLPDIHIVETI 120
 QY 124 LLSVISLNEPNTFSPANVDASVMFRKWRDSKGD--KEYAEIIRK 167
 DB 121 MISVISMLADPNQSDPANVDAA---KEWREDNRNGEKKRVARCVRK 163
 RESULT 27
 AAY27371
 ID AAY27371 standard; Protein: 170 AA.
 AC AAY27371;
 DT 19-NOV-1999 (first entry)
 XX
 XX Human ubiquitin-conjugating enzyme (HUCE-1).
 XX
 KW Human: ubiquitin-conjugating enzyme; HUCE-1; neurodegenerative disorder;
 KW cancer; immune disorder; AIDS; Alzheimer's disease; amnesia; allergy;
 KW anemia; asthma; AIDS; atherosclerosis; bronchitis; cholecystitis;
 KW Crohn's disease; inflammation; rheumatoid arthritis; scleroderma.
 XX
 CS Homo sapiens.
 XX
 PN US5952161-A.
 XX
 PD 14-SEP-1999.
 XX
 PF 26-JAN-1999; 99US-0237507.
 XX
 PR 21-AUG-1997; 97US-0918723.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Corley NC, Lal P;
 XX
 DR WPI: 1999-560362/47.
 DR N-PSDB: AAX91112.
 XX
 XX Detection of ubiquitin conjugating enzyme polynucleotides useful for
 PT the diagnosis of cancers, immune disorders or microbial infections -
 XX
 PS Disclosure: Fig 1A-E; 30pp; English.
 XX
 CC The invention provides a human ubiquitin-conjugating enzyme (HUCE-1) and
 CC methods for detecting polynucleotides (PNS) which encode HUCE-1. The
 CC detection of HUCE-1 PNS can be used for the diagnosis of conditions or
 CC disorders which are associated with expression of HUCE-1. Such conditions
 CC include e.g. cancers, neurodegenerative disorders, such as akathisia,

CC Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar
 CC disorder, catatonia, cerebral neoplasms, depression, Down's
 CC syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease,
 CC multiple sclerosis, neurofibromatosis, Parkinson's disease, paraneoplastic
 CC psychoses, schizophrenia, and Tourette's disorder, and immune disorders
 CC such as AIDS, Addison's disease, adult respiratory distress syndrome,
 CC allergies, anemia, asthma, atherosclerosis, bronchitis, cholecystitis,
 CC Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis,
 CC diabetes mellitus, emphysema, erythema nodosum, atrophic gastritis,
 CC glomerulonephritis, gout, Graves' disease, hyperesinophilia, irritable
 CC bowel syndrome, lupus erythematosus, multiple sclerosis, myasthenia
 CC gravis, myocardial or pericardial inflammation, polymyositis, rheumatoid
 CC arthritis, osteoarthritis, osteoporosis, pancreatitis, scleroderma,
 CC Sjogren's syndrome, and autoimmune thyroiditis, complications of cancer,
 CC hemodialysis, extracorporeal circulation, viral, bacterial, fungal,
 CC parasitic, protozoal and helminthic infections and trauma. Such assays
 CC may also be used to evaluate the efficacy of a particular therapeutic
 CC treatment regimen in animal studies, in clinical trials, or in monitoring
 CC the treatment of an individual patient. The present sequence represents
 CC the HUCE-1 polypeptide.
 XX
 SQ Sequence 170 AA:
 Query Match 30.58; Score 460.5; DB 20; Length 170;
 Best Local Similarity 52.48; Pred. No. 3.7e-39;
 Matches 87; Conservative 28; Mismatches 44; Indels 7; Gaps 3;
 QY 6 MISSOKALM:--ELKSLOEPPVEGFRITLVDSDLYNNEVAIFGLPNTLYEGGYFKAHIK 63
 DB 1 MTEIQSALLRLRQIAELKNKPVGFSAGLIDDDLYNNEV:ITGPPDTLYEGGVFKAHLT 60
 QY 64 FPIDYPSPTFRFLTKMWHENIYENGDCVCSILHPPVDDPQSGELPSEWNPQWRTI 123
 DB 61 FKDYPLRPKMKFTETWHPNVDKNSDVCISILHEPGEDKYGYEKPEERWLPDIHIVETI 120
 QY 124 LLSVISLNEPNTFSPANVDASVMFRKWRDSKGD--KEYAEIIRK 167
 DB 121 MISVISMLADPNQSDPANVDAA---KEWREDNRNGEKKRVARCVRK 163
 NFSULT 18
 AAW92251
 ID AAW92251 standard; Protein: 170 AA.
 AC AAW92251;
 DT 01-APR-1999 (first entry)
 XX
 XX Human HUCE-1 protein.
 XX
 DE HUCE-1; ubiquitin-conjugating enzyme; human; treatment; disorder;
 KW cancer; immune disorder; neurodegenerative disorder; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN US5863779-A.
 XX
 PD 26-JAN-1999.
 XX
 PF 21-AUG-1997; 97US-0918723.
 XX
 PR 21-AUG-1997; 97US-0918723.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Corley NC, Lal P;
 XX
 DR WPI: 1999-131307/11.
 DR N-PSDB: AAV74301.
 XX
 XX DNA encoding ubiquitin-conjugating enzyme - useful for producing
 PT recombinant enzyme
 XX

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136352.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137572.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140359.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143442.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144425.
PR 19-JUL-1999; 99US-0144431.
PR 19-JUL-1999; 99US-0144432.
PR 19-JUL-1999; 99US-0144433.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144532.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151065.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

```

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 28-OCT-1999; 99US-0162142.

Query Match 29.3% Score 452; Da 2; Length 167;
Best Local Similarity 59.3%; Pred No. 2, 3e-38;
Matches 85; Conservative 33; Mismatches 37; Indels 12; Gaps 3;

QY 6 MTSSKALME--ELKSLQEFVEGPRITIVDESILYNNVEAIFESPTLYEGGYKAIK 63
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MANNQASILLQKQKULCKKPFVDSAGIVDEKNVFOKSVSIMPDPDLYRGFFENATMS 60

QY 64 FFLDYVSPPTFRFTKMKHPNTIYNGVCISILHPVDPQSGSELPSSRNPTQNVRTI 123
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 FPNFVSPPTVTFSEMHHPNTYSDGKVCISILHPGCDPHGYELASERWTPVHIVKSI 120

QY 124 LLSVLSILNEPNTSPANDASVMFKWRDCKGKQKQYARILRKQVS 170
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 VLSITSLNSPNDSPANVEA--KEWDRRAE-----FRKKVS 157

RESULT 21
AAG54106
ID AAG54106 standard; Protein: 167 AA.
XX AAG54106;
AC AAG54106;
D: 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SHQ ID NO: 68953.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130892.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

```

XX	PN	EP1033405-A2.	9905-0121825.
XX	PN		9905-0123180.
XX	PD	06-SEP-2000.	9905-0123548.
XX	PD		9905-0125788.
XX	PD		9905-0126264.
XX	PD	25-FEB-2000; 2000EP-0301439.	9905-0126785.
XX	PD		9905-0127462.
XX	PD	25-FEB-1999;	9905-0128234.
XX	PD	05-MAR-1999;	9905-0128714.
XX	PD	09-MAR-1999;	9905-0129845.
XX	PD	23-MAR-1999;	9905-0130077.
XX	PD	25-MAR-1999;	9905-0130449.
XX	PD	29-MAR-1999;	9905-0130510.
XX	PD	01-APR-1999;	9905-0130891.
XX	PD	06-APR-1999;	9905-0131449.
XX	PD	08-APR-1999;	9905-0132048.
XX	PD	16-APR-1999;	9905-0132407.
XX	PD	19-APR-1999;	9905-0132484.
XX	PD	21-APR-1999;	9905-0132485.
XX	PD	23-APR-1999;	9905-0132486.
XX	PD	23-APR-1999;	9905-0132487.
XX	PD	28-APR-1999;	9905-0132863.
XX	PD	30-APR-1999;	9905-0134256.
XX	PD	04-MAY-1999;	9905-0134256.
XX	PD	05-MAY-1999;	9905-0134256.
XX	PD	06-MAY-1999;	9905-0134256.
XX	PD	07-MAY-1999;	9905-0134256.
XX	PD	07-MAY-1999;	9905-0134256.
XX	PD	11-MAY-1999;	9905-0134256.
XX	PD	14-MAY-1999;	9905-0134256.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134231.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136352.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137562.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141642.
PR 01-JUL-1999; 99US-0141654.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

```

PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 26-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match      29.98; Score 452; DB 21; Length 175;
Restrict local; Similarity 50.98; Pred. No. 2.9e-38;
Matches 65; Conservative 33; Mismatches 37; Indels 12; Gaps 3;

Qy 6 MTSQKALML-ELKIQEEPEVEGRITIVDESGLYNNNEVAIFGLPNTLYEGGYFKAHIK 63
Db 9 MANKQSLLIQLEKLUCKKIVQFNSAGLVKNNVQFQSVSIMGVPTLYEGGFNAMS 68

Qy 64 FPDYPSPTFRFLTKMMHNIYENDVCISILHPVDPPQSGELPSRWNPNTONVRI 123
Db 69 FPEYVPSPPTVTFTSENNHNPVYSDKVCISILHPGDDPHGYELASERWTPVHIVESI 128

Qy 124 LISVISLINEPNTSPANVDASVWFKRWKSDKCKDKEVAETIRKQVS 170
Db 129 VISTSHLSGPNDESANVEAA---KEWRDNRAE-----FRKKVS 165

RESULT 23
AAG54105
ID AAG54105 standard; Protein: 175 AA.
XX
AC AAG54105;
XX
XX
DI 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein; fragment SEQ ID NO: 68952.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW Hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX
XX EPJ33405-A2.
XX
XX 06-SEP-2000.
XX
XX 25 FEB-2000; 200SEP-0301439.
XX
XX 25 FEB-1999; 990S-0121825.
XX 05-MAR-1999; 990S-0123180.
XX 09-MAR-1999; 990S-0123948.
XX 23 MAR-1999; 990S-0125788.
XX 25-MAR-1999; 990S-0126264.
XX 29-MAR-1999; 990S-0126765.
XX 01-APR-1999; 990S-0127462.
XX 06-APR-1999; 990S-0128234.
XX 08-APR-1999; 990S-0128714.
XX 16-APR-1999; 990S-0129845.
XX 19-APR-1999; 990S-0130077.
XX 21-APR-1999; 990S-0130449.
XX 23-APR-1999; 990S-0130510.
XX 28 APR-1999; 990S-0130891.
XX 28 APR-1999; 990S-0131449.
XX 30-APR-1999; 990S-0132048.
XX 30-APR-1999; 990S-0132407.
XX 04-MAY-1999; 990S-0132484.
XX 05-MAY-1999; 990S-0132485.
XX 06-MAY-1999; 990S-0132486.
XX 06-MAY-1999; 990S-0132487.
XX 07-MAY-1999; 990S-0132863.
XX 11-MAY-1999; 990S-0134256.

PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136362.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139115.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 02-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 05-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.

```


PR 11-MAY-1999; 990S-0134356.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136521.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0135782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139452.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142320.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.

PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 18-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.

```

PR 22-OCT-1999; 99US-016098;
PR 22-OCT-1999; 99US-016098;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

Query Match      29.8%; Score 450; DB 21; Length 166;
Best Local Similarity 52.1%; Pred. No. 4.3e-38;
Matches 85; Conservative 30; Mismatches 42; Indels 6; Gaps 3;

QY 8 SSQKALMI--HLKSIQHPVEGFRILLVDESDLYNNEVAIFGLPNTILYGGYFKAHKEFP 65
DB 2 ASQASILLQKQKDLCKHPVDGFSACLVDEKNIFEWSTLIGPDILYRGGFNAIMTFP 61
QY 66 IDYPSPPTRELTXXHNPINVENGDVCLSTLHPVDGSGELPSEKWNPTQNVRTILL 125
DB 62 QNTPSPPTVRFTSDMHHPNVYSDGRVCISILHPPGDDPSGYELASFRWTPVHTVESLMI 121
QY 126 SVISLNEPNTFSPANVDASVMEKWRDCKGK-DKEYAFIIRK 167
DB 122 SIIISLSPNDSPANVEA---KEMRKDRDFKKYSRCVRK 161

RESULT 25
AAG39542
ID AAG39542 standard; Protein: 166 AA.
AC AAG39542;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48941.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression; control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405 A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825;
XX 05-MAR-1999; 99US-0121800;
XX 09-MAR-1999; 99US-0123548;
XX 23-MAR-1999; 99US-0125788;
XX 25-MAR-1999; 99US-0126264;
XX 29-MAR-1999; 99US-0126785;
XX 01-APR-1999; 99US-0127462;
XX 06-APR-1999; 99US-0128234;
XX 08-APR-1999; 99US-0128714;
XX 16-APR-1999; 99US-0129845;
XX 19-APR-1999; 99US-0130077;
XX 21-APR-1999; 99US-0130449;
XX 23-APR-1999; 99US-0130510;
XX 28-APR-1999; 99US-0130691;
XX 30-APR-1999; 99US-0131449;
XX 30-APR-1999; 99US-0132048;
XX 04-MAY-1999; 99US-0132407;
XX 05-MAY-1999; 99US-0132484;
XX 06-MAY-1999; 99US-0132485;
XX 06-MAY-1999; 99US-0132486;
XX 06-MAY-1999; 99US-0132487;
XX 07-MAY-1999; 99US-0132863;
XX 11-MAY-1999; 99US-0134256;
XX 14-MAY-1999; 99US-0134218;
XX 14-MAY-1999; 99US-0134219;
XX 14-MAY-1999; 99US-0134221;
XX 14-MAY-1999; 99US-0134370;
XX 18-MAY-1999; 99US-0134768;
XX 19-MAY-1999; 99US-0134941;
XX 20-MAY-1999; 99US-0135124;
XX 21-MAY-1999; 99US-0135353;
XX 24-MAY-1999; 99US-0135629;
XX 25-MAY-1999; 99US-0136021;
XX 27-MAY-1999; 99US-0136392;
XX 28-MAY-1999; 99US-0136782;
XX 01-JUN-1999; 99US-0137222;
XX 03-JUN-1999; 99US-0137528;
XX 04-JUN-1999; 99US-0137502;
XX 07-JUN-1999; 99US-0137724;
XX 08-JUN-1999; 99US-0138094;
XX 10-JUN-1999; 99US-0138540;
XX 10-JUN-1999; 99US-0138847;
XX 14-JUN-1999; 99US-0139119;
XX 16-JUN-1999; 99US-0139452;
XX 16-JUN-1999; 99US-0139453;
XX 17-JUN-1999; 99US-0139492;
XX 18-JUN-1999; 99US-0139454;
XX 18-JUN-1999; 99US-0139455;
XX 18-JUN-1999; 99US-0139456;
XX 18-JUN-1999; 99US-0139457;
XX 18-JUN-1999; 99US-0139458;
XX 18-JUN-1999; 99US-0139459;
XX 18-JUN-1999; 99US-0139460;
XX 18-JUN-1999; 99US-0139461;
XX 18-JUN-1999; 99US-0139462;
XX 18-JUN-1999; 99US-0139463;
XX 18-JUN-1999; 99US-0139750;
XX 18-JUN-1999; 99US-0139763;
XX 21-JUN-1999; 99US-0139817;
XX 22-JUN-1999; 99US-0139899;
XX 23-JUN-1999; 99US-0140153;
XX 23-JUN-1999; 99US-0140354;
XX 24-JUN-1999; 99US-0140695;
XX 28-JUN-1999; 99US-0140823;
XX 29-JUN-1999; 99US-0140991;
XX 30-JUN-1999; 99US-0141287;
XX 01-JUL-1999; 99US-0141842;
XX 01-JUL-1999; 99US-0142154;
XX 02-JUL-1999; 99US-0142055;
XX 06-JUL-1999; 99US-0142390;
XX 08-JUL-1999; 99US-0142803;
XX 09-JUL-1999; 99US-0142920;
XX 12-JUL-1999; 99US-0142977;
XX 13-JUL-1999; 99US-0143542;
XX 14-JUL-1999; 99US-0143624;
XX 15-JUL-1999; 99US-0144005;
XX 16-JUL-1999; 99US-0144085;
XX 16-JUL-1999; 99US-0144086;
XX 19-JUL-1999; 99US-0144325;
XX 19-JUL-1999; 99US-0144331;
XX 19-JUL-1999; 99US-0144332;
XX 19-JUL-1999; 99US-0144333;
XX 19-JUL-1999; 99US-0144334;
XX 19-JUL-1999; 99US-0144335;
XX 20-JUL-1999; 99US-0144352;
XX 20-JUL-1999; 99US-0144632;
XX 20-JUL-1999; 99US-0144884;
XX 21-JUL-1999; 99US-0144814;
XX 21-JUL-1999; 99US-0145086;
XX 21-JUL-1999; 99US-0145088;
XX 22-JUL-1999; 99US-0145085;
XX 22-JUL-1999; 99US-0145087;
XX 22-JUL-1999; 99US-0145089;
XX 22-JUL-1999; 99US-0145192;

```

PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	29.8%	Score 450;	DB 21;	Length 166;
Best local Similarity	52.1%	Pred. No. 4.3e-38;		
Matches	85;	Conservative	30;	Mismatches 42; Indels 6; Gaps

QY	8	SSQAKML--ELASQGEPEVGEFRITILVDESOLYKWEVAIFGLPNTLYEGGYKHAHKKP	65
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	2	ASQASLLQKQKDKCKHPVQGSAGIVDEKNTFFWSTIIIGPPDLYEGGFNAHFFP	61
QY	66	HDYPSPTFRFLTKWHPNTLYENGDCVCSILHPDVPDQSGELPSPRNWPTQNVRIIL	125
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	62	QNYNPSPTVRFSTDMWHPNYSQGRVCISILHPGDDPSQGYELASERWTPVHTVESIML	121
QY	126	SVISILNEPNTIFSPANVASVNFKKWRDQSGK--DKYEALRK	167
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	122	SIISHGSGNDGESPANLEAA---KEWRDKHDPFKKKVSGRCVRK	16;

RESULT 26
AAG39541
ID AAG39541 standard: Proteir.; 189 AA.
XX
XX AC
XX AAG39541;
18-OCT-2000 (first entry)
XX
XX D-
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 48940.

Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS	Arabidopsis thaliana.
PN	
XX	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 9905-0121825.
PR	05-MAR-1999; 9905-0123180.
PR	09-MAR-1999; 9905-0123548.
PR	23-MAR-1999; 9905-0125788.
PR	23-MAR-1999; 9905-0126264.
PR	29-MAR-1999; 9905-0126785.
PR	01-APR-1999; 9905-0127462.
PR	06-APR-1999; 9905-0128234.
PR	08-APR-1999; 9905-0128714.
PR	16-APR-1999; 9905-0129845.
PR	19-APR-1999; 9905-0130077.
PR	21-APR-1999; 9905-0130449.
PR	23-APR-1999; 9905-0130510.
PR	23-APR-1999; 9905-0130891.
PR	28-APR-1999; 9905-0131449.
PR	30-APR-1999; 9905-0132048.
PR	30-APR-1999; 9905-0132407.
PR	04-MAY-1999; 9905-0132484.
PR	05-MAY-1999; 9905-0132485.
PR	06-MAY-1999; 9905-0132486.

PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136322.
PR 27-MAY-1999; 990S-0136382.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 01-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137622.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140595.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140591.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142255.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.

PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 22-SEP-1999; 990S-0154779.
PR 23-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.

PR	21-OCT-1999;	99US-0163815.	PR	06-MAY-1999;	99US-0132487.
PR	22-OCT-1999;	99US-0160980.	PR	07-MAY-1999;	99US-0132487.
PR	22-OCT-1999;	99US-0160980.	PR	07-MAY-1999;	99US-0132487.
PR	22-OCT-1999;	99US-0160989.	PR	11-MAY-1999;	99US-0134256.
PR	25-OCT-1999;	99US-0161404.	PR	14-MAY-1999;	99US-0134218.
PR	25-OCT-1999;	99US-0161405.	PR	14-MAY-1999;	99US-0134219.
PR	25-OCT-1999;	99US-0161406.	PR	14-MAY-1999;	99US-0134221.
PR	26-OCT-1999;	99US-0161359.	PR	14-MAY-1999;	99US-0134270.
PR	26-OCT-1999;	99US-0161360.	PR	14-MAY-1999;	99US-0134370.
PR	26-OCT-1999;	99US-0161361.	PR	18-MAY-1999;	99US-0134768.
PR	28-OCT-1999;	99US-0161520.	PR	19-MAY-1999;	99US-0134941.
PR	28-OCT-1999;	99US-0161920.	PR	20-MAY-1999;	99US-0135124.
PR	28-OCT-1999;	99US-0161993.	PR	21-MAY-1999;	99US-0135353.
PR	29-OCT-1999;	99US-0162142.	PR	24-MAY-1999;	99US-0135629.
PR	29-OCT-1999;	99US-0162142.	PR	25-MAY-1999;	99US-0136021.
PR	29-OCT-1999;	99US-0162142.	PR	27-MAY-1999;	99US-0136392.
PR	29-OCT-1999;	99US-0162142.	PR	28-MAY-1999;	99US-0136782.
PR	29-OCT-1999;	99US-0162142.	PR	01-JUN-1999;	99US-0137222.
PR	29-OCT-1999;	99US-0162142.	PR	03-JUN-1999;	99US-0137528.
PR	29-OCT-1999;	99US-0162142.	PR	04-JUN-1999;	99US-0137550.
PR	29-OCT-1999;	99US-0162142.	PR	07-JUN-1999;	99US-0137724.
PR	29-OCT-1999;	99US-0162142.	PR	08-JUN-1999;	99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138540.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138647.
PR	29-OCT-1999;	99US-0162142.	PR	14-JUN-1999;	99US-0139119.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999;	99US-0139452.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999;	99US-0139453.
PR	29-OCT-1999;	99US-0162142.	PR	17-JUN-1999;	99US-0139492.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139454.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139455.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139456.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139457.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139458.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139459.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139460.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139461.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139462.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139463.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999;	99US-0139464.
PR	29-OCT-1999;	99US-0162142.	PR	22-JUN-1999;	99US-0139817.
PR	29-OCT-1999;	99US-0162142.	PR	22-JUN-1999;	99US-0139899.
PR	29-OCT-1999;	99US-0162142.	PR	23-JUN-1999;	99US-0140353.
PR	29-OCT-1999;	99US-0162142.	PR	23-JUN-1999;	99US-0140354.
PR	29-OCT-1999;	99US-0162142.	PR	24-JUN-1999;	99US-0140695.
PR	29-OCT-1999;	99US-0162142.	PR	28-JUN-1999;	99US-0140823.
PR	29-OCT-1999;	99US-0162142.	PR	29-JUN-1999;	99US-0140991.
PR	29-OCT-1999;	99US-0162142.	PR	30-JUN-1999;	99US-0141287.
PR	29-OCT-1999;	99US-0162142.	PR	01-JUL-1999;	99US-0141842.
PR	29-OCT-1999;	99US-0162142.	PR	01-JUL-1999;	99US-0142154.
PR	29-OCT-1999;	99US-0162142.	PR	02-JUL-1999;	99US-0142055.
PR	29-OCT-1999;	99US-0162142.	PR	06-JUL-1999;	99US-0142390.
PR	29-OCT-1999;	99US-0162142.	PR	08-JUL-1999;	99US-0142803.
PR	29-OCT-1999;	99US-01621			

[illegible]

PR 30-APR-1999; 99US-0132407;
PR 04-MAY-1999; 99US-0132484;
PR 05-MAY-1999; 99US-0132485;
PR 06-MAY-1999; 99US-0132486;
PR 06-MAY-1999; 99US-0132487;
PR 07-MAY-1999; 99US-0132463;
PR 11-MAY-1999; 99US-0134256;
PR 14-MAY-1999; 99US-0134218;
PR 14-MAY-1999; 99US-0134219;
PR 14-MAY-1999; 99US-0134221;
PR 14-MAY-1999; 99US-0134370;
PR 18-MAY-1999; 99US-0134768;
PR 19-MAY-1999; 99US-0134941;
PR 20-MAY-1999; 99US-0135124;
PR 21-MAY-1999; 99US-0135353;
PR 24-MAY-1999; 99US-0135629;
PR 25-MAY-1999; 99US-0136021;
PR 27-MAY-1999; 99US-0136352;
PR 28-MAY-1999; 99US-0136782;
PR 01-JUN-1999; 99US-0137222;
PR 03-JUN-1999; 99US-0137528;
PR 04-JUN-1999; 99US-0137502;
PR 07-JUN-1999; 99US-0137724;
PR 08-JUN-1999; 99US-0138094;
PR 10-JUN-1999; 99US-0138540;
PR 10-JUN-1999; 99US-0138847;
PR 14-JUN-1999; 99US-0139119;
PR 16-JUN-1999; 99US-0139452;
PR 16-JUN-1999; 99US-0139453;
PR 17-JUN-1999; 99US-0139492;
PR 18-JUN-1999; 99US-0139454;
PR 18-JUN-1999; 99US-0139455;
PR 18-JUN-1999; 99US-0139456;
PR 18-JUN-1999; 99US-0139457;
PR 18-JUN-1999; 99US-0139458;
PR 18-JUN-1999; 99US-0139459;
PR 18-JUN-1999; 99US-0139460;
PR 18-JUN-1999; 99US-0139461;
PR 18-JUN-1999; 99US-0139462;
PR 18-JUN-1999; 99US-0139463;
PR 18-JUN-1999; 99US-0139750;
PR 18-JUN-1999; 99US-0139763;
PR 21-JUN-1999; 99US-0139827;
PR 22-JUN-1999; 99US-0139899;
PR 23-JUN-1999; 99US-0140352;
PR 23-JUN-1999; 99US-0140354;
PR 24-JUN-1999; 99US-0140655;
PR 26-JUN-1999; 99US-0140823;
PR 29-JUN-1999; 99US-0140891;
PR 30-JUN-1999; 99US-0141287;
PR 01-JUL-1999; 99US-0141842;
PR 01-JUL-1999; 99US-0142154;
PR 02-JUL-1999; 99US-0142055;
PR 06-JUL-1999; 99US-0142380;
PR 08-JUL-1999; 99US-0142603;
PR 09-JUL-1999; 99US-0142920;
PR 12-JUL-1999; 99US-0142977;
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PR 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PR 16-JUL-1999; 99US-0144086;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144632;
PR 20-JUL-1999; 99US-0144684;
PR 21-JUL-1999; 99US-0144414;
PR 21-JUL-1999; 99US-0145086;
PR 21-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 06-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 18-OCT-1999; 99US-0159638;
PR 21-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;

PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160766.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161942.
 PR 28-OCT-1999; 99US-0161943.
 PR 29-OCT-1999; 99US-0162142.

Query Match 29.18; Score 439; DB 21; Length 166;
 Best Local Similarity 50.98; Pred. No. 5.9e-37;
 Matches 83; Conservative 31; Mismatches 43; Indels 6; Gaps 3;

QY 9 SSQKALML--ELKSQLEPVEGFRITLVDSLYNWEVAIFGLPNTLYEGGYKAKIKP 65
 DB 2 NSQCHLLQKLDCKIPVLFGRSAGVDEKNFEWSVILGPDITLYEGGFFVAINMSP 61

QY 56 IDPYSPTPTFRFLTKMHPNIVENGDCVCSILHPPVDDPOSGELPSFRWNPIONVRIIL 125
 DB 62 QNVSPPTVRETSDIHFNVPDGRVCSILHPPGDDPSGYELASERWIPVITVESIML 121

QY 126 SVLSLNEPNIISPANDASVYFRRKRDGSKGDKKEYAEIIRK 167
 DB 122 SIIISM:SGINDSPANVEAA--KENREKKDEKKKVSCVVK 161

RESULT 31:
 ASB71776
 AC AAW71776;
 XX
 DT 26-MAR-2002 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO 42120.
 KW Drosophila melanogaster polypeptide SEQ ID NO 42120.
 KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
 XX Drosophila melanogaster.
 PN W0200:71042-A2.
 XX
 PD 27-SEP-2001.
 PR 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 XX
 PT Venter JC, Adams M, et al. PWO, Myers EW;
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL15879.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions
 XX
 PS Disclosure; SEQ ID NO 42120; 21pp - Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent.
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (AB501840-ABL16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 168 AA;

Query Match 28.84; Score 434.5; DB 22; Length 168;
 Best Local Similarity 49.78; Pred. No. 1.7e-36;
 Matches 87; Conservative 23; Mismatches 54; Indels 11; Gaps 2;

QY 6 MTSSQKALML--ELKSQLEPVEGFRITLVDSLYNWEVAIFGLPNTLYEGGYKAKIK 63
 DB 1 MSRIQASLIILNROLSLQRPVVEGFSAGLVSDSDIKKWFVYIGPDDTLYEGGFFKAHLJ 60

QY 64 FPDYPSPTPTFRFLTKMHPNIVENGDCVCSILHPPVDDPOSGELPSFRWNPIONVRIIL 123
 DB 61 FPKYPPRPYKMKFTEIWHPNIDKAGDVCSILHIEPGDDKXGWEKAEERWLPVITVEII 120

QY 124 LLSVISLNEPNIISPANDASVYFRRKRDGSKGDKKEYAEIIRKQVSKATKAEK 178
 DB 121 LLSVISLNEPNIISPANDASVYFRRKRDGSKGDKKEYAEIIRKQVSKATKAEK 166

RESULT 32
 AAG44743
 ID AAG44743 standard; Protein; 140 AA.
 XX
 AC AAG44743;
 XX
 DT 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 56080.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 05-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 99US-0132486.


```

PR 21-OCT-1999; 99US-0160625;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

Query Match 27.4%; Score 414; DH 21; Length 140;
Best Local Similarity 54.3%; Pred. No. 1.7e-34;
Matches 75; Conservative 25; Mismatches 36; Indels 2; Gaps 1;

QY 6 WTSSQKALM--ELKSQEPVEGRKILVDSDLYNNVAIFGLPNILYEGYFKAHIK 63
DB 1 MANNGASILLQKLIKCKKPVDSAGLVDFKNVQKSVSIMGPPDTLYEGGFNAIMS 60

QY 64 FPIIDYPSPTFERLIKMMHPIYENGDCVCSILHPPVCDPQSGELPSEKRNPTONVTI 123
DB 61 FPIENYSPPTVTFISEMHPIYSDGKVC:SLHPPGDPPHGYELASERWTPVHIVESI 120

QY 124 ILISVISLLNEPNTFSAN 14;
DB 121 VLSTISMLSGPNDXPAN 138

QY
DB
DE
AC AAG44742;
ID AAG44742 standard; Protein; 164 AA.
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 56079.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 200CEP-0301439.
XX
XX 25-FEB-1999; 99US-0121825;
XX 05-MAR-1999; 99US-0121800;
XX 09-MAR-1999; 99US-0123548;
XX 23-MAR-1999; 99US-0125788;
XX 25-MAR-1999; 99US-0126264;
XX 29-MAR-1999; 99US-0126785;
XX 01-APR-1999; 99US-0127462;
XX 06-APR-1999; 99US-0128234;
XX 08-APR-1999; 99US-0128714;
XX 16-APR-1999; 99US-0129845;
XX 19-APR-1999; 99US-0130077;
XX 21-APR-1999; 99US-0130449;
XX 23-APR-1999; 99US-0130510;
XX 23-APR-1999; 99US-0130891;
XX 28-APR-1999; 99US-0131449;
XX 30-APR-1999; 99US-0132048;
XX 30-APR-1999; 99US-0132407;
XX 04-MAY-1999; 99US-0132484;
XX 05-MAY-1999; 99US-0132485;

PR 06-MAY-1999; 99US-0132486;
PR 06-MAY-1999; 99US-0132487;
PR 07-MAY-1999; 99US-0132863;
PR 11-MAY-1999; 99US-0134256;
PR 14-MAY-1999; 99US-0134218;
PR 14-MAY-1999; 99US-0134219;
PR 14-MAY-1999; 99US-0134221;
PR 14-MAY-1999; 99US-0134370;
PR 18-MAY-1999; 99US-0134768;
PR 19-MAY-1999; 99US-0134941;
PR 20-MAY-1999; 99US-0135124;
PR 21-MAY-1999; 99US-0135353;
PR 24-MAY-1999; 99US-0135629;
PR 25-MAY-1999; 99US-0136021;
PR 27-MAY-1999; 99US-0136392;
PR 28-MAY-1999; 99US-0136782;
PR 01-JUN-1999; 99US-0137222;
PR 03-JUN-1999; 99US-0137528;
PR 04-JUN-1999; 99US-0137502;
PR 07-JUN-1999; 99US-0137724;
PR 08-JUN-1999; 99US-0138094;
PR 10-JUN-1999; 99US-0138540;
PR 10-JUN-1999; 99US-0138847;
PR 14-JUN-1999; 99US-0139119;
PR 16-JUN-1999; 99US-0139452;
PR 16-JUN-1999; 99US-0139453;
PR 17-JUN-1999; 99US-0139492;
PR 18-JUN-1999; 99US-0139454;
PR 18-JUN-1999; 99US-0139455;
PR 18-JUN-1999; 99US-0139456;
PR 18-JUN-1999; 99US-0139457;
PR 18-JUN-1999; 99US-0139458;
PR 18-JUN-1999; 99US-0139459;
PR 18-JUN-1999; 99US-0139460;
PR 18-JUN-1999; 99US-0139461;
PR 18-JUN-1999; 99US-0139462;
PR 18-JUN-1999; 99US-0139463;
PR 18-JUN-1999; 99US-0139750;
PR 18-JUN-1999; 99US-0139763;
PR 21-JUN-1999; 99US-0139817;
PR 22-JUN-1999; 99US-0139899;
PR 23-JUN-1999; 99US-0140353;
PR 23-JUN-1999; 99US-0140354;
PR 24-JUN-1999; 99US-0140695;
PR 28-JUN-1999; 99US-0140823;
PR 29-JUN-1999; 99US-0140991;
PR 30-JUN-1999; 99US-0141287;
PR 01-JUL-1999; 99US-0141842;
PR 01-JUL-1999; 99US-0142154;
PR 02-JUL-1999; 99US-0142055;
PR 06-JUL-1999; 99US-0142390;
PR 08-JUL-1999; 99US-0142803;
PR 09-JUL-1999; 99US-0142920;
PR 12-JUL-1999; 99US-0142977;
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PR 15-JUL-1999; 99US-0144005;
PR 16-JUL-1999; 99US-0144085;
PR 16-JUL-1999; 99US-0144086;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144632;
PR 20-JUL-1999; 99US-0144684;
PR 21-JUL-1999; 99US-0144814;
PR 21-JUL-1999; 99US-0145086;
PR 21-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145085;
PR 22-JUL-1999; 99US-0145087;

```

PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147302.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147303.
 PR 09-AUG-1999; 99US-0147453.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149729.
 PR 23-AUG-1999; 99US-0149702.
 PR 23-AUG-1999; 99US-0149730.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152163.
 PR 13-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157177.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158469.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 27.4%; Score 414; DB 21; Length 164;
 Best local Similarity 54.3%; Pred. No. 2.2e-34;
 Matches 75; Conservative 25; Mismatches 36; Indels 2; Gaps 1;
 QY 6 MTSSQKALML--ELKSLOEPEVEGFRITLVDESLYNWEVAIFGLPNTLYEGGYKAHIK 63
 DB 25 MANNQASLLLOKQKDLCKKPYDGSAGLVDERKNVFORNSVINGPPDLYEGGFENAIMS 84
 QY 64 FPDYPSPTFFRLTKMHPNIYNCDCISIIHPPVDDPQSGELPSEKRNWPTQNVRTI 123
 DB 85 FPNYPSPTFTFTSEMHPNYSKGVCSILHPPGDPHGYELASERWTPVHTVESI 144
 QY 124 LLSVISLLNEPNTFSPAN 141
 DB 145 VLSTISMLSGPNDEXPAN 162

RESULT 34
 ABB60648
 ID ABB60648 standard; Protein; 167 AA.
 XX
 AC ABB60648;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 8736.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WC200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PE 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-061415O.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL04751.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 8736; Zipp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB57737-ABH72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/publised_pct_sequences.
XX
SQ Sequence 167 AA:
Query Match 25.7%; Score 389; DB 22; Length 167;
Best Local Similarity 47.9%; Pred. No. 6.6e-32;
Matches 79; Conservative 23; Mismatches 57; Indels 6; Gaps 1;
QY 8 SSQALMLELKLQEPVFGFRITLLVDSGLYNNEVAIFGLPMLVGGYKAKIKFPID 67
Db 4 SALRLMAEYKQLTLDPPEGVAGPISDNFFEWELIAGPCTCEGGVFPARLIPTD 63
QY 68 YPSPPTFRFLTKMHPNIYENGVCISLHPVDVDPQSGELPSEKRNPTQVTRILLSV 127
Db 64 YPLSPKMKFTCDMFHFNIFADGRVCISLHAPGDDPMGYELSAEHWSPQSVKILLSV 123
QY 128 ISLINEPTSPANVDASVMFRKWRDSKDKKEYAEIIRKQVSAT 172
Db 124 VSMIAFNEQESGANVDAAMWREQSD-----EFNAIARLRVKT 162
RESULT 35
AAV01786
ID AAV01786 standard; Protein: 165 AA.
XX
AC AAV01786;
XX
DE 21-DEC-1999 (first entry)
XX
DE Human ubiquitin conjugating enzyme 7 (HUBC7).
XX
KW Ubiquitin conjugating enzyme 7; HUBC7; UBC7; human;
KW protein degradation; cystic fibrosis; therapy.
XX
OS Homo sapiens.
XX
FI Key Location/Qualifiers
FI Thiolester-bond 89 /note: "thioester formation with ubiquitin"
XX
PN W09950421 A1.
XX
PD 07-OCT-1999.
XX
PF 23-MAR-1999; 99WC-0300919.
XX
PR 27-MAR-1998; 96GB-0006490.
PR 09 APR-1998; 96GB-0007533.
XX
PA (UYLE-) UNIV LEEDS.
XX
PI Markham AF, Robinson PA;
XX
DR WPI: 1999-531322/50.
DR N-PSDB: AA220068, AA220069.
XX
PT Novel polypeptides used to treat clinical conditions resulting from
PT ubiquitin conjugating enzyme 7, UBC7, mediated protein degradation
XX
PS Claim 1(a): Fig 3: 37pp; English.
XX
PS This sequence represents novel human ubiquitin conjugating enzyme 7
CC (HUBC7), as deduced from isolated cDNA (see AA220068) and genomic
CC DNA (see AA220069) clones. The invention provides HUBC7
CC polynucleotides, protein and peptides, modified HUBC7 proteins that
CC have a dominant negative effect on HUBC7 activity, a means of

CC manufacturing recombinant HUBC7, prokaryotic and eukaryotic host
CC cells, vectors, and antibodies raised against HUBC7, a method for
CC identifying and purifying HUBC7. Also provided are: a method for
CC the identification and/or purification of proteins or fragments of
CC proteins that are capable of interacting with HUBC7, e.g. human
CC UBC6 homologues or analogues; a method for screening molecules that
CC interact with HUBC7 so as to affect functional activity; an assay
CC kit for the determination of HUBC7 activity in vitro; a method of
CC monitoring degradation of cystic fibrosis transmembrane conductance
CC regulator protein or mutated variants in an ELISA using antibodies
CC produced to HUBC7; and a pharmaceutical composition comprising a
CC HUBC7 nucleic acid or protein used to treat clinical conditions
CC resulting from HUBC7 mediated protein degradation, e.g. cystic
XX fibrosis.
XX
SQ Sequence 165 AA:
Query Match 25.2%; Score 380.5; DB 20; Length 165;
Best Local Similarity 50.3%; Pred. No. 6.4e-31;
Matches 80; Conservative 18; Mismatches 54; Indels 7; Gaps 2;
QY 11 KALMLELKSIOFEPVEGFRITLVDSGLYNNEVAIFGLPMLVGGYKAKIKFPID 70
Db 7 KRLMAEYKQLTLNPPGIVAGPMNEENFFEWELIAGPCTCEGGVFPAILSPDLYPL 66
QY 71 SPPTFRFLTKMHPNIYENGVCISLHPVDVDPQSGELPSEKRNPTQVTRILLSV 130
Db 67 SPPKWRFTCEMFHFNIPDGRVCISLHAPGDDPMGYELSAEHWSPQSVKILLSV 126
QY 131 LNEPNTSPANVDASVMFRKWRDSKDKKEYAEIIRKQV 169
Db 127 LAEPNDESGANVDASKM---WRD-----DREQFYKIAKQI 158
RESULT 36
AAV01786
ID AAV01786 standard; Protein: 165 AA.
XX
AC AAV01786;
XX
DE 25-JUN-1999 (first entry)
XX
DE Human ubiquitin-conjugating enzyme HUBI-2.
XX
KW Human: ubiquitin-conjugating enzyme; HUBI-2; HUBI; cell cycle;
KW cell signaling; antagonist; neoplastic disorder;
KW UCS dependent proteolysis; neuronal disorder; immune disorder;
KW developmental disorder.
XX
OS Homo sapiens.
XX
PN W09915659 A2.
XX
PD 01-APR-1999.
XX
PF 22-SEP-1998; 98WO-0519970.
XX
PR 06-NOV-1997; 97US-0965689.
PR 23-SEP-1997; 97US-0933750.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Corley NC, Hillman JL, Lal P;
XX
DR WPI: 1999-244422/20.
DR N-PSDB: AAX34870.
XX
PT New isolated human vesicle trafficking protein
XX
PS Claim 1: Fig 2A-D; 70pp; English.
XX
PS The present sequence represents a human ubiquitin-conjugating enzyme
CC HUBI-2. HUBI is expressed in neoplastic, immune, developmental and

neurological disorders where HUB1 plays a role in the cell cycle and in cell signaling. Antagonists of HUB1 can be used to prevent or treat neoplastic disorders. Since HUB1 appears to be involved in UCS dependent proteolysis and is found in neuronal tissues, an antagonist of HUB1 may be administered to a subject to prevent or treat a neuronal disorder, e.g. akathisia, Alzheimer's disease, amnesia, epilepsy, Huntington disease, multiple sclerosis, Parkinson's disease, etc... An antagonist of HUB1 may be administered to a subject to prevent or treat an immune disorder, e.g. AIDS, Addison's disease, adult respiratory distress syndrome, allergies, asthma, etc... Antagonists can also be administered to a subject to prevent or treat a developmental disorder such as renal tubular acidosis, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, etc...

XX Sequence 165 AA;

Query Match 25.2%; Score 380.5; DR 23; Length 165;
Best Local Similarity 50.3%; Pred. No. 6.4e-31;
Matches 80; Conservative 18; Mismatches 54; Indels 7; Gaps 2;
QY 11 KALMELKSLQEPVAGFRITLVDSIDLYNWEVAIFGLPNTLYEGGYFKAHKKFPIDYPPY 70
DB 7 KRLMAEYKQLTLNPPGIVAGPNEENFEWEALIMGPEDTOFFGVFPALISFPLDYPL 66
QY 71 SPPTFRLTKMHHPNIYENGVCISILHPPVDDQSGSELPSEKWNPTQNVRTILLSVISL 130
DB 67 SPKMKRFCTCFMHNIPGRCVCSILHAPGDDPMGYESSAERWSPQSVKILLSVYSVM 126
QY 131 LNFNPTSPANVDASVMFRKWRDSKGRKAEIRKOV 169
DB 127 LAEPNDESGANVDASKM---WRD---DREGFYKIAKQI 158

RESULT 37
AAY84866
ID AAY84865 standard; Protein: 165 AA.

XX AC AAY84865;

XX DT 21-AUG-2000 (first entry)

XX DE Amino acid sequence of ubiquitin-conjugating enzyme HSUBC18.

XX KW Ubiquitin-conjugating enzyme; HSUBC14; HSUBC15; HSUBC18; proteasome; endoplasmic reticulum; protein degradation; mutant protein maturation; deltax508; cystic fibrosis transmembrane conductance regulator; cystic fibrosis; alpha1-AT; emphysema; liver disease.

XX OS Homo sapiens.

XX PN WO200023599-A2.

XX PZ 27-APR-2000.

XX PF 21-OCT-1999; 99WO-US24563.

XX PR 21-OCT-1998; 98US-0105064.

XX PA (LEUK-) LEUKOSITE INC.

XX PI Chau V.

XX PS WPI: 2000-339698/29.

XX DR N-PSDB; AAA14977.

XX PT New - Purified ubiquitin-conjugating enzymes HSUBC14, HSUBC15, and HSUBC18 useful for treating proteasome-mediated endoplasmic reticulum protein degradation defects, e.g. cystic fibrosis, emphysema, liver disease.

XX PS Claim 32; Fig 3; 71pp; English.

XX CC The present sequence represents a human ubiquitin-conjugating enzyme.

CC designated HSUBC18. The specification also describes HSUBC14 and HSUBC15. The ubiquitin-conjugating enzymes allow proteasome-mediated endoplasmic reticulum protein degradation and promote the maturation of mutant proteins such as deltax508 and alpha1-AT. The ubiquitin-conjugating enzymes are useful for treating diseases associated with proteasome-mediated endoplasmic reticulum protein degradation, cystic fibrosis caused by failure of deltax508 or cystic fibrosis transmembrane conductance regulator (CFTR) precursors to mature into functional CFTR, emphysema caused by failure of mutant alpha1-AT to be secreted, and liver disease caused by an accumulation of mutant alpha1-AT in the endoplasmic reticulum. Their binding molecules and nucleic acids specifically complementary or homologous to nucleic acids encoding these enzymes may be used in conventional assays to detect the presence, absence or quantity of HSUBC14, HSUBC15, and HSUBC18 in a sample.

XX Sequence 165 AA;

Query Match 25.2%; Score 380.5; DB 21; Length 165;
Best Local Similarity 50.3%; Pred. No. 6.4e-31;
Matches 80; Conservative 18; Mismatches 54; Indels 7; Gaps 2;
QY 11 KALMELKSLQEPVAGFRITLVDSIDLYNWEVAIFGLPNTLYEGGYFKAHKKFPIDYPPY 70
DB 7 KRLMAEYKQLTLNPPGIVAGPNEENFEWEALIMGPEDTOFFGVFPALISFPLDYPL 66
QY 71 SPPTFRLTKMHHPNIYENGVCISILHPPVDDQSGSELPSEKWNPTQNVRTILLSVISL 130
DB 67 SPKMKRFCTCFMHNIPGRCVCSILHAPGDDPMGYESSAERWSPQSVKILLSVYSVM 126
QY 131 LNFNPTSPANVDASVMFRKWRDSKGRKAEIRKOV 169
DB 127 LAEPNDESGANVDASKM---WRD---DREGFYKIAKQI 158

RESULT 38

AAG70896

ID AAG70896 standard; Protein: 167 AA.

XX AC AAG70896;

XX DT 27-JUL-2001 (first entry)

XX DE C albicans apoptosis associated protein #76.

XX KW Yeast; fungus; apoptosis; infection; proliferative disease; vaccine; autoimmune disease; ischaemia; neurodegeneration.

XX CS Candida albicans.

XX PN WO200102550-A2.

XX PZ 11-JAN-2001.

XX PF 03-JUL-2000; 2000WO-BE00077.

XX PR 01-JUL-1999; 99EP-0870141.

XX PA (JANC) JANSSEN PHARM NV.

XX PI Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;

XX PS Nelissen BJM, Reekmans RJ;

XX DR WPI: 2001-367042/38.

XX DR N-PSDB; AAH29932.

XX PT Yeast and fungal nucleic acids encoding proteins involved in a pathway leading to programmed cell death, useful for treating proliferative disorders, yeast and fungal infections, or for preventing apoptosis in certain diseases.

XX PS Claim 24; Fig 2; 218pp; English.

XX CC The present invention provides the protein and coding sequences of a


```

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

Query Match 24.08; Score 362; DB 21; Length 125;
  Hest Local Similarity 54.08; Prod. No. 3.4e-29;
Matches 67; Conservative 21; Mismatches 26; Indels 10; Gaps

QY 47 GLPNTLYEGGYKAIKIPIDYPSPTFRFLTMMHPIYENGDCVCSILHPPVDDPQS 106
Db 2 GPPDTLYEGGFFNIMSPENYPYSPPTVTFTSEMHPIYSDGKYCSILHPPGDPHG 61
QY 107 GELSEHWNPNTONVITLLSVISLNEPNTFSPANVDASVMFKWHDGKDKYAEIIR 166
Db 62 VELASERWPTVHTVESIVLSISMLSGPNDESPANVEA---KEWRCRAE-----FR 111

QY 167 KQVS 170
Db 112 MKVS 115

RESULT 40
AAC54107
ID AAC54107 standard; Protein; 125 AA.
XX
AC AC
XX
DT DT
XX
18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SWQ ID NO: 68954.
XX
KW protein identification; signal transduction pathway; metabolic pathwa
KW hybridisation assay; genetic mapping; gene expression control; promot
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD
XX
06-SEP-2000.
XX
25-FEB-2000; 2000EP-0301439.
XX
25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

```

PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134270.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138340.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 19-JUN-1999; 99US-0139461.
 PR 19-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140595.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142350.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143342.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144352.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145143.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149358.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149910.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 10-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.

```

PR 25-OCT-1999: 990S-0161404.
PR 25-OCT-1999: 990S-0161405.
PR 25-OCT-1999: 990S-0161406.
PR 26-OCT-1999: 990S-0161359.
PR 26-OCT-1999: 990S-0161360.
PR 26-OCT-1999: 990S-0161361.
PR 28-OCT-1999: 990S-0161320.
PR 28-OCT-1999: 990S-0161592.
PR 28-OCT-1999: 990S-0161593.
PR 29-OCT-1999: 990S-0162142.

Query Match      24.00; Score 362; DB 21; Length 125;
Rest Local Similarity 34.00; Pred. No. 3.4e-29;
Matches 67; Conservative 21; Mismatches 26; Indels 10; Gaps 2;

QY 47 C.P.N.I.Y.R.C.C.Y.K.A.H.I.K.F.I.D.Y.P.S.P.P.I.F.R.F.I.T.K.M.H.H.N.I.Y.E.N.C.D.V.C.I.S.I.L.H.P.P.V.D.D.Q.S 106
Db 2 G.P.D.T.L.Y.E.G.G.F.N.A.I.M.S.F.E.N.Y.P.V.S.P.T.V.T.T.S.E.M.H.E.N.Y.S.D.G.K.V.C.I.S.I.L.H.P.P.G.D.D.P.H.G 61

QY 107 G.E.L.P.S.E.R.A.N.P.T.O.N.V.R.T.I.L.L.S.V.I.L.N.E.P.N.T.E.S.P.A.N.V.D.A.S.V.N.F.K.W.R.D.S.K.G.K.D.K.E.Y.A.E.I.I.R 166
Db 62 Y.E.L.A.S.H.R.W.T.P.V.H.I.V.S.I.V.I.S.I.S.M.S.G.P.N.D.S.P.A.N.V.E.A.A.-----K.E.N.R.E.N.K.A.E.-----F.R 11;

QY 167 K.Q.V.S 170
Db 112 K.K.V.S 115

```

Search completed: April 10, 2003, 10:32:17
 Job time : 53.7694 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Computer Ltd.

QM protein - protein search, using sw model

Run on: Apr 10, 2003, 10:27:54 ; Search time 13.9341 Seconds
(without alignments)

1945.578 Million cell updates/sec

Title: US-09-930-026-2

Perfected score: 1511

Sequence: 1 MAQQQTSSQRALEELKSL.....MAAPQKPIHSGWNTSSC 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues 283224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 200%

Listing first 135 summaries

Database :

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1010	60.9	238	2 A49630	ubiquitin conjugat
2	452	29.9	167	2 S46656	ubiquitin-protein
3	450	29.8	166	2 S71205	ubiquitin-protein
4	444	29.4	169	2 T02943	ubiquitin-protein
5	444	29.4	169	2 T01329	ubiquitin-conjugat
6	439	29.1	166	2 T45703	ubiquitin-protein
7	432.5	28.6	295	2 A41241	ubiquitin-conjugat
8	428.5	28.4	164	2 A40982	ubiquitin-protein
9	425.5	28.2	213	2 S19158	ubiquitin-protein
10	417	27.6	167	2 T39286	probable ubiquitin
11	411.5	27.2	215	1 U0XFA5	ubiquitin-protein
12	404.5	26.8	168	2 A41547	ubiquitin-conjugat
13	367	24.3	166	2 T43235	ubiquitin-conjugat
14	358.5	23.7	165	2 S28951	ubiquitin-conjugat
15	337.5	22.3	237	2 T27470	hypothetical prote
16	320	20.5	172	2 A21906	ubiquitin-conjugat
17	295	19.5	152	2 T51913	ubiquitin-conjugat
18	295	19.5	152	2 A42416	ubiquitin carrier
19	295	19.5	152	2 B41222	ubiquitin-protein
20	286	18.9	152	2 A41222	ubiquitin-protein
21	285	18.9	205	2 T32959	hypothetical prote
22	284	18.8	151	2 S22529	ubiquitin-conjugat
23	283	18.7	151	2 T51931	hypothetical prote
24	281	18.6	151	2 S71430	DNA repair protein
25	280	18.5	152	2 S43783	ubiquitin-conjugat
26	279	18.5	151	2 T45220	ubiquitin-protein
27	276	18.3	152	2 S43781	ubiquitin-conjugat
28	271	17.9	151	2 A39492	RA66 DNA-repair ho
29	267.5	17.7	176	2 T40902	ubiquitin conjugat

30	266	17.6	150	2 S43782	ubiquitin-conjugat
31	265.5	17.6	156	2 S67248	hypothetical prote
32	261	17.3	155	2 T37559	ubiquitin-conjugat
33	260.5	17.2	157	2 S62571	probable ubiquitin
34	252	16.7	148	2 S32674	ubiquitin-protein
35	251.5	16.6	147	2 T39202	ubiquitin-conjugat
36	250	16.5	148	2 S32672	ubiquitin-protein
37	248	16.4	146	2 D96666	ubiquitin-conjugat
38	240.5	15.9	148	2 S22857	ubiquitin-protein
39	240.5	15.9	151	2 S35793	ubiquitin-protein
40	240.5	15.9	215	1 S12493	ubiquitin-protein
41	240	15.9	157	2 T18512	ubiquitin-conjugat
42	237	15.7	147	2 A48145	ubiquitin-conjugat
43	237	15.7	147	2 S19157	ubiquitin-conjugat
44	237	15.7	148	2 F84543	probable ubiquitin
45	235.5	15.6	158	2 JC6056	ubiquitin-protein
46	234.5	15.5	148	2 S61417	ubiquitin-protein
47	234	15.5	147	2 T23820	hypothetical prote
48	233	15.4	154	2 E86484	hypothetical prote
49	232	15.4	154	2 JC6163	ubiquitin-conjugat
50	231.5	15.3	148	2 S22858	ubiquitin-protein
51	231	15.3	147	2 F90082	ubiquitin-conjugat
52	231	15.3	147	2 F90118	ubiquitin-conjugat
53	231	15.3	147	2 F90137	ubiquitin-conjugat
54	231	15.3	147	2 D90102	ubiquitin-conjugat
55	231	15.3	147	2 H90116	ubiquitin-conjugat
56	229	15.2	147	2 T39300	ubiquitin-conjugat
57	229	15.2	218	2 B53516	ubiquitin-protein
58	227.5	15.1	118	2 S32673	ubiquitin-protein
59	227.5	15.1	166	2 T29929	hypothetical prote
60	226	15.0	147	2 S53358	ubiquitin-conjugat
61	224.5	14.9	159	2 JC5970	nuclear ubiquitin-
62	223.5	14.8	152	2 JC4894	ubiquitin-protein
63	223.5	14.8	153	2 S58092	hypothetical prote
64	223.5	14.8	195	2 D96541	hypothetical prote
65	223	14.8	133	2 T54552	hypothetical serin
66	223	14.8	147	2 S53359	ubiquitin-conjugat
67	222.5	14.7	153	2 C86304	probable ubiquitin
68	222	14.7	147	2 T59365	ubiquitin-conjugat
69	219	14.5	162	2 T16646	hypothetical prote
70	215.5	14.3	146	2 T37532	ubiquitin-conjugat
71	213.5	14.1	154	2 T13578	ubiquitin-conjugat
72	209	13.8	163	2 B96818	ubiquitin-protein
73	207	13.7	157	2 S52414	ubiquitin-protein
74	205.5	13.6	183	2 S52661	ubiquitin-protein
75	205.5	13.6	185	2 S43785	ubiquitin-protein
76	202.5	13.4	147	2 D90126	ubiquitin-conjugat
77	200	13.2	194	2 S57619	ubiquitin-conjugat
78	196	13.0	247	2 B42856	ubiquitin carrier
79	195.5	12.9	144	2 G90107	ubiquitin-conjugat
80	194	12.8	170	2 A96663	hypothetical prote
81	193.5	12.8	200	2 A40757	ubiquitin-conjugat
82	193	12.8	180	2 T24059	hypothetical prote
83	192.5	12.7	189	2 S43784	ubiquitin-protein
84	192	12.7	152	2 T27167	hypothetical prote
85	191.5	12.7	126	2 T50603	ubiquitin-conjugat
86	191	12.6	187	2 S43786	ubiquitin-protein
87	189.5	12.5	235	2 T48741	probable ubiquitin
88	189	12.5	199	2 T08465	ubiquitin-protein
89	188.5	12.5	161	2 S39483	ubiquitin-conjugat
90	188.5	12.5	199	2 T33629	hypothetical prote
91	187	12.4	183	2 A3516	ubiquitin-protein
92	187	12.4	183	2 JC4308	ubiquitin-protein
93	187	12.4	184	2 A34506	23K ubiquitin carr
94	186.5	12.3	198	2 T40123	ubiquitin-conjugat
95	186	12.3	204	2 A53848	ubiquitin-conjugat
96	184.5	12.2	160	2 T46009	E2 ubiquitin-conju
97	184	12.2	184	2 T50342	ubiquitin-conjugat
98	179	11.8	139	2 T14451	ubiquitin-conjugat
99	178.5	11.8	227	2 T37499	probable ubiquitin
100	177.5	11.7	161	2 A96785	putative ubiquitin
101	177.5	11.7	177	2 T00789	ubiquitin-protein
102	174	11.5	309	2 T34195	hypothetical prote

A:Molecule type: DNA
 A:Residues: 1-166 <VNA>
 A:Cross-references: EMBL:U33757; MID:992703; PID:4992704
 R:van Nocker, S.; Walker, J.M.; Vierstra, R.D.
 J. Biol. Chem. 271, 12150-12158, 1996
 A:Title: The Arabidopsis thaliana UBC7/13/14 genes encode a family of multiubiquitin chain
 A:Reference number: 221063; MUID:8647807; PMID:8647807
 A:Accession: T51253
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-166 <VNA>
 A:Cross-references: EMBL:U33757; PID:AAAC43321.1
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Gene: ubc7
 A:Introns: 15/1; 48/2; 62/1; 115/3; 141/3
 C:Superfamily: human ubiquitin-protein ligase E2
 C:Keywords: ligase

Query Match 29.8%; Score 450; DB 2; Length 166;
 Best Local Similarity 52.2%; Pred. No. 4.4e-26;
 Matches 85; Conservative 30; Mismatches 42; Indels 6; Gaps 3;

QY 8 SSOKALML--ELKSLQEPVEGRITLVDSOLYNEVAIFGLPNTLYEGGYFKAHKE 65
 DB 2 ASQASILLQKQLKOLAKHPVDSAGISVDEKNIFFWSTVITIGPDTLYEGGFNAIMTF 61

QY 66 IDVPYSPPTFRELTKMHPNIYENGDCVCSILHPPVDDPSGSGELPSERNWNTONVRIIL 125
 DB 62 QNPNSPPSVRTSEMHPNYPVDSGRCVCSILHPPGDPNGYELASERWTPVHTVIESIV 123

QY 126 SVISILNENPTFSPANVDASVMFKKRSKCK-DKEYAEIRK 167
 DB 124 LSTISLSSPNDESANIEA---KWKREDEKKKVRQVRK 164

RESULT 4
 T02943
 ubiquitin-conjugating enzyme E2 - maize
 N:Alternate names: ubiquitin carrier protein
 C:Species: Zea mays (maize)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Jun-2000
 C:Accession: T02943
 R:de Vries, A.; Cordts, S.; Dresselhaus, I.
 Plant Physiol. 118, 1161, 1998
 A:Title: Molecular Characterization of a cDNA Encoding an Ubiquitin Carrier Protein (UBC
 A:Reference number: 214787; MUID:95105148; PMID:9890811
 A:Accession: T02943
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-169 <DEV>
 A:Cross-references: EMBL:AJ002959; PID:CAA05772.2
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 29.4%; Score 444; DB 2; Length 169;
 Best Local Similarity 49.4%; Pred. No. 1.4e-28;
 Matches 81; Conservative 37; Mismatches 40; Indels 6; Gaps 3;

QY 7 TSSOKALML--ELKSLQEPVEGRITLVDSOLYNEVAIFGLPNTLYEGGYFKAHKE 64
 DB 4 TIIQASILLQKQLKOLAKHPVDSAGISVDSNFWQVITIGPDTLYEGGFNAIMTF 63

QY 65 IDVPYSPPTFRELTKMHPNIYENGDCVCSILHPPVDDPSGSGELPSERNWNTONVRIIL 124
 DB 64 QNPNSPPSVRTSEMHPNYPVDSGRCVCSILHPPGDPNGYELASERWTPVHTVIESIV 123

QY 125 SVISILNENPTFSPANVDASVMFKKRSKCK-DKEYAEIRK 167
 DB 124 LSTISLSSPNDESANIEA---KWKREDEKKKVRQVRK 164

ubiquitin-conjugating enzyme E2 - maize
 C:Species: Zea mays (maize)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Nov-1999
 C:Accession: T01329
 R:Morlier, J.M.; Gregerson, R.S.
 Plant Physiol. 116, 1191, 1998
 A:Title: Isolation and DNA sequence analysis of an aspartate aminotransferase cDNA c1
 A:Reference number: 214296
 A:Accession: T01329
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-169 <MOR>
 A:Cross-references: EMBL:AF032468; MID:g2641618; PID:AAAC12662.1; PID:g2641619
 A:Experimental source: strain Ze
 C:Genetics:
 A:Gene: ubc7
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 29.4%; Score 444; DB 2; Length 169;
 Best Local Similarity 49.4%; Pred. No. 1.4e-28;
 Matches 81; Conservative 37; Mismatches 40; Indels 6; Gaps 3;

QY 7 TSSOKALML--ELKSLQEPVEGRITLVDSOLYNEVAIFGLPNTLYEGGYFKAHKE 64
 DB 4 SASQASILLQKQLKOLAKHPVDSAGISVDSNFWQVITIGPDTLYEGGFNAIMTF 63

QY 65 IDVPYSPPTFRELTKMHPNIYENGDCVCSILHPPVDDPSGSGELPSERNWNTONVRIIL 124
 DB 64 QNPNSPPSVRTSEMHPNYPVDSGRCVCSILHPPGDPNGYELASERWTPVHTVIESIV 123

QY 125 SVISILNENPTFSPANVDASVMFKKRSKCK-DKEYAEIRK 167
 DB 124 LSTISLSSPNDESANIEA---KWKREDEKKKVRQVRK 164

RESULT 6
 T45703
 ubiquitin-protein ligase (EC 6.3.2.19) UBC13 [similarity] - Arabidopsis thaliana
 N:Alternate names: protein F18L15.180; ubiquitin conjugating enzyme, E2
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 03-Jun-2002
 C:Accession: T45703; T51254
 R:Choine, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattivello, L.; Artiguenave, F.;
 submitted to the Protein Sequence Database, November 1999
 A:Reference number: 223010
 A:Accession: T45703
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-166 <CHO>
 A:Cross-references: EMBL:ALJ33298; PID:CAE62037.2
 A:Experimental source: cultivar Columbia; BAC clone F18L15
 R:van Nocker, S.; Walker, J.M.; Vierstra, R.D.
 J. Biol. Chem. 271, 12150-12158, 1996
 A:Title: The Arabidopsis thaliana UBC7/13/14 genes encode a family of multiubiquitin
 A:Reference number: 221063; MUID:8647807; PMID:8647807
 A:Accession: T51254
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-166 <VAN>
 A:Cross-references: EMBL:U33758; PID:AAAC49322.1
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Gene: UBC13; F18L15.180
 A:Map position: 3
 A:Introns: 15/1; 49/2; 82/1; 115/3; 141/3
 A:Note: F18L15.180
 C:Superfamily: human ubiquitin-protein ligase E2
 C:Keywords: ligase

Query Match 29.1%; Score 439; DB 2; Length 166;
 Best Local Similarity 50.9%; Pred. No. 3.3e-28;
 Matches 83; Conservative 31; Mismatches 43; Indels 6; Gaps 3;


```

Best Local Similarity 42.8%; Score 417; DB 2: Length 167;
Matches 95; Conservative 33; Mismatches 77; Indels 17; Gaps 5;

QY 13 IMLEKLSQEEVGEPRITLLVDESGLYNNWEVAIFGLPNTLYEGGYEKAHKIKPEIDYKXP 72
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 6 LLAERYKNLIVNFSHFKIS-VNEDNITENDVILKGPDDLYEGGLFKAKIVPPKYPYRP 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 73 PTPRELLIKWHPNINYEKDCISILHPPVDDPQSGELPSEKWNPTQNVRTLLSVLSILN 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 PKITFTSEMMHPNINYPDGRICISILH---GNAAEQGTMWSPAQKIDTLLSVLSILN 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 133 EPNTFSPANVDASVWFKK---HUSKGDKEYAEIIRKOVSAIKAPAEKIGKVPITLAE 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 120 EPNPDSHPANVDAAKSYKRY-LYKEDLESYPMEYKKTIVKKSIDECSEADIEYFKKNPVNV- 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 190 YCIKTKVPNSNNSLLDGLYDUIIDDEDEEDADUCYDD 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 ---LTVSPDYDYDEMELOTYIILVYDDEDEEDDEE- MDE 213
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
T39286
probable ubiquitin-protein ligase (EC 6.3.2.19) e2 - fission yeast (Schizosaccharomyces
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Jun-2002
C:Accession: T39286
R:McDougal, R.C.; Rajadream, M.A.; Barrell, B.G.; Ramsperger, U.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, July 1999
A:Reference number: 221841
A:Accession: T39286
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1 167 <MCD>
A:Cross-references: PM3:AL096851; PIDN:CAB50972.1; GSPDB:GN00067; SPDB:SPBC1105.09
A:Experimental source: strain: 972h; cosmid c105
C:Genetics:
A:Gene: SPDB:SPBC1105.09
A:Map position: 2
A:Features: 5/4; 63/3; 73/3; 149/1
C:Superfamily: human ubiquitin-protein ligase E2
C:Keywords: ligase

Query Match 27.6%; Score 417; DB 2: Length 167;
Best Local Similarity 46.8%; Pred. No. 2e-26;
Matches 76; Conservative 33; Mismatches 48; Indels 6; Gaps 1;

QY 7 ISSQKALMELKS-QEEFVEGFRITLLVDESGLYNNWEVAIFGLPNTLYEGGYEKAHKIKPEI 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 4 SASQQLAKKQKLEIQKPPQGSFVGLVDKSIPEFVMIQPHDLYEGGFPHATISPPQ 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 67 DYDYSPTFRELTKWHPNINYEKDCISILHPPVDDPQSGELPSEKWNPTQNVRTLLS 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 64 DYPLAPPKMKETIELHFNHPNGEVCISILHPPGDDKGYGEGADAGWLPVHSPTILIS 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 127 VLSILNEPTVSPANVDASVWFKKWDKSGDKKEYAEIIRKQV 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 124 VLSMLSPNDESPANIDAAKEFRE-----NPQEFKKVRRLV 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
UQKFS
ubiquitin-protein ligase (EC 6.3.2.19) E2 - African swine fever virus (strain BA71V)
C:Species: African swine fever virus; ASFV
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 03-Jun-2002
C:Accession: F39448
R:Rodriguez, J.M.; Salas, M.L.; Vinuela, E.
Virology 186, 40-52, 1992
A:Title: Genes homologous to ubiquitin-conjugating proteins and eukaryotic transcription
A:Reference number: A19448; MUID:92687485; PMID:1309282
A:Accession: F39448
A:Molecule type: DNA
A:Residues: 1-215 <KOD>
A:Cross-references: GB:M77121; NID:9210618; PIDN:AAA42704.1; PID:9210624

```

C:Superfamily: African swine fever virus ubiquitin-protein ligase E2
C:Keywords: ligase; protein degradation

Query Match 27.2%; Score 411.5; DB 1: Length 215;
Best Local Similarity 41.7%; Pred. No. 7.6e-26;
Matches 91; Conservative 40; Mismatches 72; Indels 15; Gaps 6;

QY 13 IMLEKLSQEEVGEFRITLLVDSGLYNNWEVAIFGLPNTLYEGGYEKAHKIKPEIDYKXP 72
:
DB 6 LIAEYRILINSENFKIS-VNENNITENDVILRPPDLYEGGLFKAKVAFPPPYAP 64
: :

QY 73 PTPRELLIKWHPNINYEKDCISILHPPVDDPQSGELPSEKWNPTQNVRTLLSVLSILN 132
:
DB 65 PKITFTSEMMHPNINYPDGRICISILH---GNAAEQGTMWSPAQKIDTLLSVLSILN 119
: :

QY 133 EPNTFSPANVDASVWFKKWDKSGDKKEYAEIIRKOVSAIKAEAKSGKVPITLAEYCI 192
:
DB 120 EPNPDSHPANVDAAKSYKRY-VYKEDLESYPMEYKKTIVKKSIDECSPEDIEYKNAASNV- 177
: :

QY 193 KTKVPSN--DNSSDILYDD----LYDDDDIDDEDEED 224
:
DB 178 -PPPSDAYEDECEDDITVILYDDDEDEEDDD 214
: :

RESULT 12
A41547
ubiquitin-conjugating enzyme E2 - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 30-Jun-1992 #sequence_revision 18-Sep-1992 #text_change 24-Nov-1999
C:Accession: A41547
R:Van Nocker, S.; Vierstra, R.D.
Proc. Natl. Acad. Sci. U.S.A. 88, 10297-10301, 1991
A:Title: Cloning and characterization of a 20-kDa ubiquitin carrier protein from whe
A:Reference number: A41547; MUID:92052257; PMID:1658801
A:Accession: A41547
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-168 <VAN>
A:Cross-references: GB:M74077
A:Note: the sequence given in Fig. 2 differs from that shown in lacking residue 16
C:Superfamily: human ubiquitin-protein ligase E2

Query Match 26.8%; Score 404.5; DB 2: Length 168;
Best Local Similarity 50.7%; Pred. No. 2e-25;
Matches 74; Conservative 28; Mismatches 37; Indels 7; Gaps 3;

QY 24 PVGFRITLLVDESGLYNNWEVAIFGLPNTLYEGGYEKAHKIKPEIDYKSPPTFRFLTKMWH 83
:
DB 23 PSMGFOLGFWDDSNVFEQVITLIGPPETLYDGYFNAINSPNYKSPPTVRFTSEMH 82
: :

QY 84 PNIVENGDCVCSILHPPVDDPQSGELPSEKWNPTQNVRTLLSVLSILNEPNTFSPANVD 143
:
DB 83 PNYPDGRVCISI-HPGDDPNGYELASERWTPVHTVESIVLSISMLSPNDESPANIE 141
: :

QY 144 ASYMFKKWDKSGDKKEYAEIIRKQV 169
:
DB 142 AA---KDWRE---KODEFKKKRRV 161
: :

RESULT 13
T43235
ubiquitin-conjugating enzyme ubcp3 - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: Ubcp3
C:Species: Schizosaccharomyces pombe
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T43235
R:Osaka, F.; Seino, H.; Seno, T.; Yamao, F.
Mol. Cell. Biol. 17, 3388-3397, 1997
A:Title: An ubiquitin-conjugating enzyme in fission yeast that is essential for the
A:Reference number: Z22356; MUID:97299689; PMID:9154838
A:Accession: T43235
A:Status: preliminary; translated from: GB/EMBL/DBJ
A:Molecule type: mRNA

A:Residues: 1-166 <OSA>
 A:Cross-references: EMBL:D85544; NID:q2190251; PIDN:BAA20373.1; PID:q2190252
 C:Genetics:
 A:Gene: ubcP3
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 24.3%; Score 367; DB 2: Length 166;
 Best Local Similarity 45.1%; Pred. No. 2 le-22;
 Matches 73; Conservative 25; Mismatches 58; Indels 6; Gaps 2;

QY 11 KALMLKLSLOEPVEGFRITLVDSLYNWEVAIFGLPNTLYEGGYFKAHKEPDIY 70

DB 8 RHIMKEYELTENPQGITAGISNEODEFTWQXLIQGMXCTPPEGSLYPAILKPFSDYPL 67

QY 71 SPPTFRFLTKMHPNIYENDVCISILHPVPDDPQSGELPSEHWNTQVRIILLVSL 130

DB 68 GPPTLKCEFFHPNFKKTVGCIISILHAPGDPPNMYESSSESWSPQSVEXILLSVSM 127

QY 131 LNEPNTFSPANVDASVNFRRKWKDGKQKAYAFIRKQVSAI 172

DB 128 LAEPNDESGANIDCKM---WRDR---EEYRVVRLARKT 163

RESULT 14
 S2895:
 ubiquitin-conjugating enzyme UbcP - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein Y9711.12; protein YMR022w; ubiquitin-conjugating enzyme QR18
 C:Species: Saccharomyces cerevisiae
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
 C:Accession: S28951; S29741; S54024; S29338
 R:Vassal, A.; Boulet, A.; Decoster, E.; Faye, G.
 Biochim. Biophys. Acta 1132, 21-213, 1992

A:Title: QR18, a novel ubiquitin-conjugating enzyme in Saccharomyces cerevisiae.

A:Reference number: S28951; MUID:94054327; EMBL:1327.48

A:Accession: S28951

A:Molecule type: DNA

A:Residues: 1-165 <AS>

A:Cross-references: EMBL:X66829; NID:q4256; PIDN:CAA47302.1; PID:q4257

R:Jungmann, J.; Reins, H.A.; Schobert, C.; Jentsch, S.

Nature 361, 369-371, 1993

A:Title: Resistance to cadmium mediated by ubiquitin-dependent proteolysis.

A:Reference number: S29741; MUID:93149278; PMID:8381213

A:Accession: S29741

A:Molecule type: DNA

A:Residues: 1-165 <JFN>

A:Cross-references: EMBL:X69100; NID:q5522; PIDN:CAA48846.1; PID:q5523

R:Lye, G.; Churcher, C.M.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54014

A:Accession: S54024

A:Molecule type: DNA

A:Residues: 1-165 <LYE>

A:Cross-references: EMBL:Z49211; NID:q798922; PID:q798933; MIPS:YMR022w

A:Experimental source: strain AH972

C:Genetics:

A:Gene: SQD:QR18; Ubc7

A:Cross-references: SGD:S0004624; MIPS:YMR022w

A:Map position: 138

C:Superfamily: human ubiquitin-protein ligase E2

Query Match 23.7%; Score 358.5; DB 2: Length 165;

Best Local Similarity 47.3%; Pred. No. 9.9e-22;

Matches 70; Conservative 26; Mismatches 49; Indels 3; Gaps 1;

QY 8 SSQKALMLKLSLOEPVEGFRITLVDSLYNWEVAIFGLPNTLYEGGYFKAHKEPDI 67

DB 4 TADKRILKELQQLKDSPPGIVAGSPSENFIWDCLIQGPDTPPYANGVFNKALIEFPKD 63

QY 68 YPKSPPTFRFLTKMHPNIYENDVCISILHPVPDDPQSGELPSEHWNTQVRIILLVSL 127

DB 64 YPLSPKPLTFPSILHPNIYENVCISILHSPGDPPNMYELAEERWSPQSVEXILLSV 123

QY 128 ISLNEPNTFSPANVDASVNFRRKWKDGK 155

DB 124 MSMLSEPTSCANLIDACIL---WRDR 148

RESULT 15

T27470

hypothetical protein Y87G2A.r - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000

C:Accession: T27470

R:White, S.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20371

A:Accession: T27470

A:Status: preliminary; translated from GB/EMBL/DDDB

A:Molecule type: DNA

A:Residues: 1-237 <WLD>

A:Cross-references: EMBL:A1110500; NID:el542114; PIDN:CAB54492.1; CESP:Y87G2A.r

A:Experimental source: Clone Y87G2A

C:Genetics:

A:Gene: CESP:Y87G2A.r

A:Introns: 15/1; 86/1; 129/3; 196/1

C:Superfamily: African swine fever virus ubiquitin-protein ligase E2

Query Match 22.3%; Score 337.5; DB 2: Length 237;

Best Local Similarity 33.0%; Pred. No. 7.8e-20;

Matches 77; Conservative 24; Mismatches 61; Indels 71; Gaps 3;

QY 11 KALMLKLSLOEPVEGFRITLVDSLYNWEVAIFGLPNTLYEGGYFKAHKEP----- 65

DB 7 KRIMTEYKELITRPPGIIAAPIDEDNFFEWELITGPETCFANGVFPARITFPQVRLQ 66

QY 66 ----- 65

DB 67 ITLPFSLTQPKLIFWRNSHIFSRNFTLKSFQVSLRNSVFSDNFESNKAIKKDCDTP 126

QY 66 --IDYPSPTFRFLTKMHPNIYENDVCISILHPVPDDPQSGELPSEHWNTQVRI 123

DB 127 FKDYPLSPKMFRTGIFHPNIYAGRVCSILHAPGDDPTGYELSNERSWPQSIKI 186

QY 124 LLSVISLLNEPNTFSPANVDASVNFRRKWKDGKQKQKAYE-IIRKQVSATKAE 175

DB 187 LLSVWSMLAEPNDPSPANVSAKM---WRDRQAQFEKIAUSILVRKILCLPASE 236

RESULT 16

A21906

ubiquitin-conjugating enzyme RAD6 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein G3447; protein YGL056w

C:Species: Saccharomyces cerevisiae

C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 21-Jul-2000

C:Accession: A21906; A29592; S64062

R:Reynolds, P.; Weber, S.; Prakash, L.

Proc. Natl. Acad. Sci. U.S.A. 82, 168-172, 1985

A:Title: RAD6 gene of Saccharomyces cerevisiae encodes a protein containing a tract

A:Reference number: A21906; MUID:85113143; PMID:3881753

A:Accession: A21906

A:Molecule type: DNA

A:Residues: 1-172 <REY>

A:Cross-references: EMBL:K02962; NID:gl72348; PIDN:AAA34952.1; PID:gl72349

R:Jentsch, S.; McGrath, J.P.; Varshavsky, A.

Nature 329, 131-134, 1987

A:Title: The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating enzyme.

A:Reference number: A29592; MUID:87315384; PMID:3306404

A:Accession: A29592

A:Molecule type: protein

A:Residues: 77-87, 'X', 89-91 <JEN>

R:Kueermann, M.; Potier, S.; Souciet, J.L.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64044

A:Accession: S64062

A:Molecule type: DNA

A:Residues: 1-172 <REY>

A:Cross-references: EMBL:272584; NID:q1322557; PID:q1322558; MIPS:YGL058W
 A:Experimental source: Strain S258C

C:Genetics:

A:Gene: S4D:RAD6; UBC2

A:Cross-references: SGD:S000326; MIPS:YGL058W

A:Map position: 71

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: nucleus; thiolester bond; ubiquitination

F:88/Cross-link: thiolester (Cys) (laterchain to Gly-76 of ubiquitin) #status experiment

Query Match 20.5%; Score 310; DB 2; Length 172;
 Best Local Similarity 30.6%; Pred. No. 8.4e-16;
 Matches 66; Conservative 46; Mismatches 58; Indels 50; Gaps 4;

QY 7 TSQKALMELKLSQEPVGFRTILVDSDLYNNEVAIFGLPNTLYEGGYPKAHKFP1 66

DE 3 IPARRLRMRDKRKLQEDPPVGVSGAPSENNIMQNAVIFGPGTFFEDGTFKLVIEFSE 61

QY 67 DYPSPPTFRFLTKMWHPNYENGVCISILHPVDDPQSGELPSEKWNPTQNVRTILLS 126

DB 62 EYFNKPPTVRFELSKMHPNHYADGSCICLDIL-----ONRWPTVDVASILTS 108

QY 127 VTSILNEPNTSPANVDASVMFRKWRDSKDKKEYAEIIRKQVSATKAEAKD 186

DB 109 IQSLIDEPNPSPANVAALILF-----KDKSQYKRVKETEVSQWSDM----- 153

QY 187 LAEYCIKTKVPSNDSLLIYDGLYDDIDDEEEDAD 226

DE 154 -----DDMDDDDDDDDDDDDDDDDD 172

RESULT 17

151913

14 kDa ubiquitin conjugating enzyme - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999

C:Accession: 151913

R:Wing, S.S.: Hanville, D.

Am. J. Physiol. 267, E30-F48, 1994

A:Title: The 14 kDa ubiquitin conjugating enzyme: structure of the rat gene and regulation

A:Reference number: 151913; MUID:94324482; PMID:8048511

A:Accession: 151913

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-152 <RES>

A:Cross-references: EMBL:J04308; NID:q476115; PID:AA860669.1; PID:q476117

C:Genetics:

A:Introns: 15/2; 42/2; 51/1; 81/1; 110/3

C:Superfamily: human ubiquitin-protein ligase E2

Query Match 19.5%; Score 295; DB 2; Length 152;

Best Local Similarity 37.6%; Pred. No. 1.1e-16;

Matches 65; Conservative 31; Mismatches 53; Indels 24; Gaps 4;

QY 7 TSQKALMELKLSQEPVGFRTILVDSDLYNNEVAIFGLPNTLYEGGYPKAHKFP1 66

DB 3 IPARRLRMRDKRKLQEDPPVGVSGAPSENNIMQNAVIFGPGTFFEDGTFKLVIEFSE 61

QY 67 DYPSPPTFRFLTKMWHPNYENGVCISILHPVDDPQSGELPSEKWNPTQNVRTILLS 126

DB 62 EYFNKPPTVRFELSKMHPNHYADGSCICLDIL-----ONRWPTVDVASILTS 108

QY 127 VTSILNEPNTSPANVDASVMFRKWRDSKDKKEYAEIIRKQVSATKAEAKD 179

DB 109 IQSLIDEPNPSPANVAALILF-----KDKSQYKRVKETEVSQWSDM 151

RESULT 18

A42416

ubiquitin carrier protein E2 - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000

C:Accession: A42416

R:Wing, S.S.: Dumas, F.; Hanville, D.

J. Biol. Chem. 267, 6495-6501, 1992

A:Title: A rabbit reticulocyte ubiquitin carrier protein that supports ubiquitin-dep

A:Reference number: A42416; MUID:92202189; PMID:1313008

A:Accession: A42416

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-152 <WIN>

A:Cross-references: GB:M62387; NID:q165779; PID:AAA31492.1; PID:q165780

A:Experimental source: reticulocyte

A>Note: sequence extracted from NCBI backbone (NCBI:89551, NCBI:89553)

C:Superfamily: human ubiquitin-protein ligase E2

Query Match 19.5%; Score 295; DB 2; Length 152;

Best Local Similarity 37.6%; Pred. No. 1.1e-16;

Matches 65; Conservative 31; Mismatches 53; Indels 24; Gaps 4;

QY 7 TSQKALMELKLSQEPVGFRTILVDSDLYNNEVAIFGLPNTLYEGGYPKAHKFP1 66

DB 3 IPARRLRMRDKRKLQEDPPVGVSGAPSENNIMQNAVIFGPGTFFEDGTFKLVIEFSE 61

QY 67 DYPSPPTFRFLTKMWHPNYENGVCISILHPVDDPQSGELPSEKWNPTQNVRTILLS 126

DB 62 EYFNKPPTVRFELSKMHPNHYADGSCICLDIL-----ONRWPTVDVASILTS 108

QY 127 VTSILNEPNTSPANVDASVMFRKWRDSKDKKEYAEIIRKQVSATKAEAKD 179

DB 109 IQSLIDEPNPSPANVAALILF-----KDKSQYKRVKETEVSQWSDM 151

RESULT 19

B41222

ubiquitin-protein ligase (EC 6.3.2.19) E2B - human

N:Alternate names: ubiquitin carrier protein E2; ubiquitin-conjugating enzyme HHR6B

C:Species: Homo sapiens (man)

C>Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 03-Jun-2002

C:Accession: B41222; S12530; S17516

R:Koken, M.H.M.; Reynolds, P.; Jaspers-Bekker, I.; Prakash, L.; Prakash, S.; Bootsma

Proc. Natl. Acad. Sci. U.S.A. 88, 8865-8869, 1991

A:Title: Structural and functional conservation of two human homologs of the yeast f

A:Reference number: A41222; MUID:92020951; PMID:1717990

A:Accession: B41222

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-152 <KOK>

A:Cross-references: GB:M74525; NID:q184045; PID:AAA35982.1; PID:q184046

R:Schneider, R.; Eckerskorn, C.; Lotzspeich, F.; Schweizer, M.

EMBO J. 9, 1431-1435, 1990

A:Title: The human ubiquitin carrier protein E2(M(r) - 17000) is homologous to the y

A:Reference number: S12530; MUID:90228340; PMID:2158443

A:Accession: S12530

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-152 <SCH>

A:Cross-references: GB:X53251; NID:q30953; PID:CAA37339.1; PID:q30954

R:Woffendin, C.; Chen, Z.; Staskus, K.; Retzel, R.F.; Piagemann, P.G.W.

Biochim. Biophys. Acta 1090, 81-85, 1991

A:Title: Mammalian mRNAs encoding protein closely related to ubiquitin-conjugating e

A:Reference number: S17516; MUID:91355233; PMID:1863845

A:Accession: S17516

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-21, 'C', '24-40, 'I', '42-53, 'R', '55-152 <WOF>

C:Genetics:

A:Gene: GDB:UBE2B; UBC2; HHR6B

A:Cross-references: GDB:128627; OMIM:174905

A:Map position: 5q23-5q31

C:Superfamily: human ubiquitin-protein ligase E2

C:Keywords: ligase; protein degradation; ubiquitination

F:88/Active site: Cys #status predicted

Query Match 19.5%; Score 295; DB 2; Length 152;

Best Local Similarity 37.6%; Pred. No. 1.1e-16;

[illegible]

RA06 DNA-repair homolog Dhrf6 - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C>Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 20-Sep-1999
C:Accession: A39392
R:Koken, M.; Reynolds, P.; Bootsma, D.; Hoelmakers, J.; Prakash, S.; Prakash, L.
Proc. Natl. Acad. Sci. U.S.A. 88, 3832-3836, 1991
A:title: Dhrf6, a *Drosophila* homolog of the yeast DNA-repair gene RAD6.
A:Reference number: A39392; MUID:91219466; PMID:1392572
A:Accession: A39392
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-151 <OK>
A:Cross-references: GB:M63791; GB:M63792; NID:g156731; PID:g156732
C:Genetics:
A:Gene: FlyBase:UbcD6
A:Cross-references: FlyBase:Fgn0004436
C:Superfamily: human ubiquitin-protein ligase E2

Query Match 17.9%; Score 271; DB 2; Length 151;
Best Local Similarity 35.6%; Pred. No. 9,7e-15;
Matches 57; Conservative 34; Mismatches 49; Indels 20; Gaps 3;

QY 7 ISSOKALMLELKSLOEPEVGEFRITLIDVESOLYKWEVAIEGLNLTIVYEGYKAHKKFPI 66
DB 3 IYAKRMRDPRKROEPPPTGVSGAPID-NNIMINWAVIGPHDITFEDGIFKLIITFTE 61
QY 67 DYPSPTTREFLTMMHNPITYENGWCISILHPVDDPQSGELPSEKRWNTQVRIILS 126
DB 62 EYPNKPPIVREKSVKFIPIVYADGGICLDIL-----QNKWSFRDVSAILTS 108
QY 127 VISLNEPNTFSPANVDASVYMKPKKRDGSKDKKEYAEIIR 166
DB 109 IQSLSDPNPNSPANSTAAQLYKENR-----REYKRVK 142

RESULT 29
T40902
ubiquitin conjugating enzyme - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T40902
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rothe, G.; Pohl, T.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z21955
A:Accession: T40902
A>Status: preliminary; translated from GB/EMBL/DDBI
A:Molecule type: DNA
A:Residues: 1-176 <WO>
A:Cross-references: EMBL:AL034564; PIDN:CAB38416.1; GSPDH:GN00068; SPDH:SHCC1259.150
A:Experimental source: strain 972h-; cosmid c1259
C:Genetics:
A:Gene: SPDH:SPCC1259.15c
A:Map position: 3
A:Introns: 32/2; 41/3; 71/3; 150/2
C:Superfamily: human ubiquitin-protein ligase E2

Query Match 17.7%; Score 267.5; DB 2; Length 176;
Best Local Similarity 39.9%; Pred. No. 2.3e-14;
Matches 57; Conservative 20; Mismatches 53; Indels 13; Gaps 1;

QY 8 SSOKALMLELKSLOEPEVGEFRITLIDVESOLYKWEVAIEGLNLTIVYEGYKAHKKFPI 67
DB 26 SVTKRLSELMSLWMSNTPGISAFPDSDNLLHWAGTIIGPSDIYYEGLKFKLSMSPAN 87
QY 68 YPSPPTTREFLTMMHNPITYENGWCISILHPVDDPQSGELPSEKRWNTQVRIILSV 127
DB 88 YPSPPTTIFTSPMHPNPVDMGSGNICLDIF-----KKKSAVYVWQIILSL 134
QY 128 ISLLNEPNTFSPANVDASVYMK 150
DB 135 QSLLEGPNNASPLNAQAELWSK 157

```

Db      1 MAVFEGGVYIKIKQELLQLLSSITSAFVVDONOLTYWVGVIITCPKDTGYSLGKFKV 60
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      61 HIKPPIIDYPSPTFRFLTKMHPNIYENGDCVCSILHPPVDDPQSGELFSRWNPQTVN 120
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      61 SLKEPQNYFPHPMKILFSPMHPNVNKGNCICLDIL-----KKNSAYNV 107
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      121 RTILLSVISLLNPNPNSPANVDASVNFRRKRDGKDKYKAIHKKQVSATKAPAE 177
      ||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      108 ETILLSLQSLGEPNRRSPLNAVAEL--W-----LADMEE-YRKVKLACYEID 154
      ||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 32
t37559
ubiquitin-conjugating enzyme e2-16 kd - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T37559
R:Wiedler, H.; Duesterhoeft, A.; Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21724
A:Accession: T37559
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-155 <MED>
A:Cross-references: EMBL:AL10509; PDB:CA854826.1; GSPDH:GN00066; SPDB:SPAC1250.03
A:Experimental source: strain 972h-; cosmid c1250
C:Genetics:
A:Gene: SPDB:SPAC1250.03
A:Map position: 1
A:Introns: 14/3; 83/3
C:Superfamily: human ubiquitin-protein ligase E2

Query Match 17.3%; Score 261; DB 2; Length 155;
Best Local Similarity 33.5%; Pred No. 6.4e-14;
Matches 54; Conservative 31; Mismatches 62; Indels 14; Gaps 3.

QY      8 SSKALMLELKLQEEPVEGRITLVDSIMYNWEVAIFGLPNILYEGGYFKAHKEPIID 67
      ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      7 SSSRLTKESDLREHPIDPIRVNLVDD-NLEHMACTALGPSDSVYAGCKFHFSLKPELD 65
      ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      68 YPSPTFRFLTKMHPNIYENGDCVCSILHPPVDDPQSGELFSRWNPQTVNRTILLSV 127
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      66 YPQPPTPIETIRIYHPNFUSEGNCVLAIKQQV-----FKDSTIKRSVLFGI 113
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      128 ISLLNPNTFSPANVDASVNFRRKRDGKCK-DKEYAEIIRK 167
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      114 LQLRHPNPDPIVASIAEQYRNDRPVDFDKIARDYVFOFAK 154
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 33
S62571
Probable ubiquitin-protein ligase (EC 6.3.2.19) hus5 - fission yeast (Schizosacchar-
N:Alternate names: protein SPAC30D11.13
C:Species: Schizosaccharomyces pombe
C:Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 03-Jun-2002
C:Accession: S62571; T38586; T43397
R:Pearson, D.; Churcher, C.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S62559
A:Accession: S62571
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-157 <PEA>
A:Cross-references: EMBL:Z67961; NID:q1065887; PDB:CAA91899.1; PID:q1065900
R:Pearson, D.; Churcher, C.M.; Barrell, H.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z21801
A:Accession: T38586
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-157 <PE2>
A:Cross-references: EMBL:Z67961; PDB:CAA91899.1; GSPDH:GN00066; SPDB:SPAC30D11.13

```


QY 9 SOKALMLLELSQEPEVEGFRITLVDESOLYNWEVAIFGLPNTLYEGGVFKAHIKFPIDY 68
 : : : ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 DB 2 ASKRITLKEKLQDKDPPTSCSAGPAVE-DMFHQAATIMGPSDPSYSGCVFLVTHFFPPDY 60

QY 69 PYSPPTRELTMMHNPNIYNGDVCSILIHPPVDPOQSGLPSEKNNPTQNVRTILLSVI 128
 : : : ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 DB 61 PKPKPVNAFTKVFHPNINSNGISCLDI2-----KFQWSPALTISKVLLSLIC 107

QY 129 SLINPEPTSPANVDASVMFR 149
 ||| :||| : : : :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 DB 108 SLATTPNDPLVPDEIAHYK 128

RESULT 35
 I39202
 ubiquitin conjugating enzyme - human
 C:Species: Homo sapiens (man)
 C>Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 08-Oct-1999
 C:Accession: I39202; S42465
 R:Scheifner, M.; Huhreque, J.M.; Howley, P.M.
 Proc. Natl. Acad. Sci. U.S.A. 91, 8797-8801, 1994
 A>Title: Identification of a human ubiquitin-conjugating enzyme that mediates the E1
 A:Reference number: I39202; MID:94377440; PMID:8090726
 A:Accession: I39202
 A:Molecule type: mRNA
 A:Residues: 1-147 <RES>
 A:Cross-references: EMBL:X78140; NID:9460809; PIDN:CAA55019.1; PID:9460810
 C:Genetics:
 A:Gene: UBCH5
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 16.6%; Score 251.5; DB 2; Length 147;
 Best Local Similarity 34.7%; Pred. No. 3.5e-13;
 Matches 52; Conservative 24; Mismatches 53; Indels 21; Gaps 3;

QY 11 KALMLELSQEPEVEGFRITLVDESOLYNWEVAIFGLPNTLYEGGVFKAHIKFPIDYPY 70
 : : : ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 DB 4 KRIOELSDLRQDPPARC-SAGPVGD-DLFHQAOATIMGPPDSAYGGVFELIVHEPTDYPF 62

QY 71 SPPTRELTMMHNPNIYNGDVCSILIHPPVDPOQSGLPSEKNNPTQNVRTILLSVLSL 130
 : : : ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 DB 63 KPCKIAFTTKIYHPNINSNGISCLDIR-----SQWSPALTISKVLLSLIC 109

QY 131 LNEPNTESPANVDASVMFRKWRDSKGDKKE 160
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 DB 110 LCDPNPDPLVPDIAQIYK-----SDKE 132

RESULT 36
 S32672
 ubiquitin-protein ligase (EC 6.3.2.19) UBC10 - Arabidopsis thaliana
 N:Alternate names: ubiquitin-conjugating enzyme UBC10
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 03-Jun-2002
 C:Accession: S32672
 R:Girol, P.; Carpenter, T.B.; van Nocker, S.; Sullivan, M.L.; Vierstra, R.D.
 submitted to the EMBL data library, August 1992
 A>Description: Homologs of the essential ubiquitin conjugating enzymes ubc1, 4, and
 A:Reference number: S32672
 A:Accession: S32672
 A:Molecule type: mRNA
 A:Residues: 1-148 <GR>
 A:Cross-references: EMBL:Z14991; NID:9297877; PIDN:CAA78715.1; PID:9297878
 C:Superfamily: human ubiquitin-protein ligase E2
 C:Keywords: ligase

Query Match 16.5%; Score 250; DB 2; Length 148;
 Best Local Similarity 35.5%; Pred. No. 4.6e-13;
 Matches 50; Conservative 25; Mismatches 52; Indels 14; Gaps 2;

QY 9 SOKALMLELSQEPEVEGFRITLVDESOLYNWEVAIFGLPNTLYEGGVFKAHIKFPIDY 68
 : : : ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 DB 2 ASKRITLKEKLQDKDPPTSCSAGPAVE-DMFHQAATIMGPSDPSYSGCVFLVTHFFPPDY 60

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45932
A:Accession: S45947
A:Molecule type: DNA
A:Residues: 1-148 <STE>
A:Cross-references: EMBL:D35951; NID:g536143; PIDN:CAA85027.1; PID:g536344; MIPS:YBR
R:Andre, B.; Cziepluch, C.; Rein, C.; Jauniaux, J.C.; Urrestarazu, A.; Vissers, S.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45993
A:Accession: S45949
A:Molecule type: DNA
A:Residues: 1-148 <ANC>
A:Cross-references: EMBL:D35951; NID:g536343; PIDN:CAA85027.1; PID:g536344; MIPS:YBR
R:van der Aart, Q.J.M.; Barthe, C.; Dolignon, F.; Aigle, M.; Crouzet, M.; Steensma, H.
Yeast 10, 959-964, 1994
A:Title: Sequence analysis of a 31 kb DNA fragment from the right arm of Saccharomyces
A:Reference number: S45462; UIDD:95076715, PMID:7985423
A:Accession: S45480
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-116, 'y', 17-148 <VAN>
A:Cross-references: EMBL:X76294
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 199
R:van der Aart, Q.J.M.
submitted to the EMBL data Library, August 1995
A:Reference number: S59702
A:Accession: S59718
A:Molecule type: DNA
A:Residues: 1-148 <VAV>
A:Cross-references: EMBL:X76294; NID:g974203; PIDN:CAAS3942.1; PID:g974207
A:Experimental source: strain S48BC
C:Genetics:
A:Gene: SGD:UHC4
A:Cross-references: SGD:S0000286; MIPS:YBR082C
A:Map position: 2R
A:Introns: 16/2
C:Superfamily: human ubiquitin-protein ligase E2
C:Keywords: blocked amino end; ligase
F:86/Active size: Cys #status predicted

	Query Match	15.9%	Score 240.5;	DB 2:	Length 148;
	Best Local Similarity	31.9%	Pred. No. 2.7e-12;		
	Matches	53;	Conservative	28;	Mismatches 54; Indels 31; Gaps 4;
QY	8	SSKALMLKSLQLSEPFVGCRITLLVDISLYNEVA:FGLPNTLYEGGYENAHIKRPID	67		
Db	2	SSSKTAKELSDLERDPPTSCSGPVGD-DIYHQASTMGPAUSYAGGVFLSLIHPTD	60		
QY	68	YPYSPPTFRFLTMMHPNIYENGDVCSILHIPPVDDPQSGLSELPNMNPNTVRLTLVS	127		
Db	61	YFPKKPKLSPTTKIYHPNINANGNICLDTH-----KDQWSPALLSLKVLLS	107		
QY	128	ISLLMEPNTPSPANVDASVMF-----RKWRDSKGCKEYA	162		
Db	108	CSSLTDANPDPIVPETAIHYTIKDRPYEAATAREM-----TKKYA	147		
RESULT	39				
S5793					
gene	endless protein - fruit fly (<i>Drosophila melanogaster</i>)				
C:Species:	<i>Drosophila melanogaster</i>				
C:Date:	20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999				
C:Accession:	S5793				
R:Thomas, J.R.					
submitted to the EMBL Data Library, June 1993					
A:Reference number:	S35793				
A:Accession:	S35793				
A>Status:	preliminary				
A:Molecule type:	mRNA				
A:Residues:	1-151 <THO>				
A:Cross-references:	GH:I20126; EMBL:D23103; NID:g304667; PID:g304668				
C:Genetics:					
A:Gene:	FlyBase:ben				

Job time : 18.2675 secs

A:Cross-references: FlyBase:FBgn000173
 C:Superfamily: human ubiquitin-protein ligase E2

Query Match 15.9%; Score 240.5; DB 2; Length 151;
 Best Local Similarity 32.2%; Pred. No. 2.9e-12;
 Matches 56; Conservative 35; Mismatches 60; Indels 23; Gaps 4;

QY 6 MISSOKALMLKLSQEPVEGFRILLYVDSLYNNWVAIFGLPNTLYEGGYFKAHKKFP 65
 DB 1 MSLPRTIKETQRLQGEPPVGIN-AIPDENNARYFHVITGPNDSPFEGGVKLEFLP 59
 QY 66 TDYPSPTFRFLTKMHHNINLVNGDVCLSI:HPPIVDPOSGGELPSFRWNPNTONVRIILL 125
 DB 60 EYPMASAKVAFIKYIYHNIDRIQRICLDV:-----KDKWSPALQIRIILL 106
 QY 126 SVISLINEPNTFSPANDVSMVKRSGKSGKDEYAEIIRKOVSAATKAEAKD 179
 DB 107 S'QALISAPNDUPLANDVAEL:--WKVNE-----APAIKNAREWTQYAYED 151

RESULT 40
 S12493
 ubiquitin-protein ligase (EC 6.3.2.19) UBC1 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YD9395.10; protein YDR177w; ubiquitin-conjugating enzyme UBC1
 C:Species: Saccharomyces cerevisiae
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: S12493; S49773
 R:Seufferl, W.; McGrath, J.P.; Jentsch, S.
 EMBJ J. 9, 4535-4541, 1990
 A:Title: UBC1 encodes a novel member of an essential subfamily of yeast ubiquitin-conjugating enzymes
 A:Reference number: S12493; MUID:91092279; PMID:2265617
 A:Accession: S12493
 A:Molecule type: DNA
 A:Residues: 1-215 <SEU>
 A:Cross-references: EMBL:X56402; NID:q4716; PIDN:CAA39812.1; PID:q4717
 R:Murphy, L.; Harris, D.E.
 submitted to the EMBL Data Library, November 1994
 A:Reference number: S49773
 A:Accession: S49773
 A:Molecule type: DNA
 A:Residues: 1-215 <MUR>
 A:Cross-references: EMBL:Z46727; NID:q1289283; PIDN:CAA86682.1; PID:g1289292; CSFDB:GN000000000
 C:Genetics:
 A:Gene: SGD:UBC1; MIPS:YDR177w
 A:Cross-references: SGD:S0002584; MIPS:YDR177w
 A:Map position: 4R
 C:Function:
 A:Description: cell cycle control; DNA repair; ligase; protein degradation
 C:Superfamily: yeast ubiquitin-protein ligase UBC1
 C:Keywords: cell cycle control; DNA repair; ligase; protein degradation

Query Match 15.9%; Score 240.5; DB 1; Length 215;
 Best Local Similarity 28.4%; Pred. No. 4.4e-12;
 Matches 62; Conservative 35; Mismatches 78; Indels 43; Gaps 5;

QY 8 SSOKALMLKLSQEPVEGFRITIVDSLYNNWVAIFGLPNTLYEGGYFKAHKKFPID 67
 DB 2 SHAKIMKEIOAVKDDPAAHITLFEVSESDIHLKGLTFLGPPGIPVEGSGEVVDIEVPME 61
 QY 68 YDPSPTFRFLTKMHHNINLVNGDVCLSI:HPPIVDPOSGGELPSFRWNPNTONVRIILL 126
 DB 62 YDFKPKMQGFTKYVHNINSSVTGATCIDIL:-----KNWSPVITLKSALIS 108
 QY 127 VISLINEPNTFSPANDVSMVKRSGKSGKDEYAEIIRKOVSAATKAEAKDGVKVPIT 186
 DB 109 LAQLIQSPENDVQ--DAEVAQHYLKDRSPFNKTAALWIKLYASR:----- 152
 QY 187 LAEYCIKTKVPSNNSSDLIYDILY--DDQ:DDDEEE 222
 DB 153 -----SNQKGNVEESDLYGIDHCLIDRESQ 179

Search completed: April 10, 2003, 10:33:45

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:39:30 : Search time 11.28 seconds
(without alignments)
1528.401 Million cell updates/sec

Title: US-09-930-026-2

Perfect score: 282
Sequence: 1 MAQVNTSSKALMLELKL.....AMAPQPKTHSGWNTHSNC 282

Scoring table: BLO62
Gapop 60.0, Gapext 60.0

Searched: 248812 seqs, 6136040 residues

Word size : 0

Total number of hits satisfying chosen parameters: 869:4

Minimum DB seq length: 3

Maximum DB seq length: 70

Post-processing: Listing first 135 summaries

Database : Published Applications_ML:
1: /cgn2_6/prodata/1/pubaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/1/pubaa/PCR_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubaa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubaa/PCRUS_PUBCOMB.pep.*
8: /cgn2_6/prodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubaa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/1/pubaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubaa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/1/pubaa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/1/pubaa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/1/pubaa/US60_PUBCOMB.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	2.8	65	10	US-09-864-761-42347
2	6	2.2	13	9	US-09-826-290-48
3	6	2.1	20	9	US-09-736-457-1855
4	6	2.1	20	9	US-09-902-943-1855
5	6	2.1	20	9	US-09-849-626-1855
6	6	2.1	20	9	US-10-017-754-1855
7	6	2.2	25	9	US-10-057-505-25
8	6	2.1	33	9	US-09-866-538-21
9	6	2.1	33	9	US-10-121-258-22
10	6	2.1	34	10	US-09-864-761-46817
11	6	2.1	34	10	US-09-864-761-48968
12	6	2.1	35	9	US-10-083-815-68
13	6	2.1	36	10	US-09-864-761-34662
14	6	2.1	37	9	US-09-575-847-5
15	6	2.1	39	10	US-09-884-681-39
16	6	2.1	45	8	US-08-424-550R-588
17	6	2.1	47	10	US-09-864-761-36758
18	6	2.1	49	10	US-09-864-761-39559
19	6	2.1	52	10	US-09-864-761-33667
20	6	2.1	52	10	US-09-864-761-41108

20	6	2.1	54	10	US-09-864-761-39215	Sequence 39215, A
21	6	2.1	54	10	US-09-864-761-44300	Sequence 44300, A
22	6	2.1	55	10	US-09-864-761-44147	Sequence 44147, A
23	6	2.1	57	10	US-09-864-761-44799	Sequence 44799, A
24	6	2.1	60	9	US-09-738-626-3524	Sequence 3524, Ap
25	6	2.1	62	10	US-09-864-761-47146	Sequence 47146, A
26	6	2.1	65	10	US-09-932-679-56	Sequence 56, Appl
27	6	2.1	67	10	US-09-912-962-33	Sequence 33, Appl
28	6	2.1	70	9	US-09-738-626-5114	Sequence 5114, Ap
29	5	1.8	6	9	US-09-727-963A-20	Sequence 20, Appl
30	5	1.8	7	9	US-09-910-346C-10	Sequence 10, Appl
31	5	1.8	7	9	US-10-083-815-2	Sequence 2, Appl
32	5	1.8	7	9	US-09-223-751A-40	Sequence 40, Appl
33	5	1.8	7	10	US-09-989-789-733	Sequence 733, Ap
34	5	1.8	7	10	US-09-989-789-1017	Sequence 1017, Ap
35	5	1.8	7	10	US-09-989-789-1222	Sequence 1222, Ap
36	5	1.8	7	10	US-09-989-789-1231	Sequence 1231, Ap
37	5	1.8	9	9	US-10-010-184A-8	Sequence 8, Appl
38	5	1.8	9	9	US-10-102-283-79	Sequence 79, Appl
39	5	1.8	9	9	US-10-102-283-176	Sequence 176, Ap
40	5	1.8	9	9	US-10-102-283-182	Sequence 182, Ap
41	5	1.8	9	10	US-09-834-765-342	Sequence 342, App
42	5	1.8	9	10	US-09-834-765-618	Sequence 618, App
43	5	1.8	9	10	US-09-919-048-79	Sequence 79, Appl
44	5	1.8	9	10	US-09-919-048-175	Sequence 176, App
45	5	1.8	9	10	US-09-919-048-182	Sequence 182, App
46	5	1.8	10	9	US-09-987-675-6	Sequence 6, Appl
47	5	1.8	10	9	US-10-102-265-16	Sequence 16, Appl
48	5	1.8	10	9	US-09-277-074-15	Sequence 15, Appl
49	5	1.8	10	10	US-09-866-824A-7	Sequence 7, Appl
50	5	1.8	10	10	US-09-834-765-211	Sequence 211, App
51	5	1.8	10	10	US-09-834-765-295	Sequence 295, App
52	5	1.8	10	10	US-09-834-765-407	Sequence 407, App
53	5	1.8	10	10	US-09-834-765-498	Sequence 498, App
54	5	1.8	10	10	US-09-834-765-503	Sequence 503, App
55	5	1.8	10	10	US-09-834-765-620	Sequence 610, App
56	5	1.8	10	10	US-09-834-765-694	Sequence 694, App
57	5	1.8	10	10	US-09-766-412-17	Sequence 17, Appl
58	5	1.8	10	10	US-09-987-655-6	Sequence 6, Appl
59	5	1.8	12	9	US-09-962-955C-17	Sequence 17, Appl
60	5	1.8	13	10	US-09-813-329-39	Sequence 39, Appl
61	5	1.8	13	10	US-09-813-329-49	Sequence 49, Appl
62	5	1.8	13	10	US-09-813-329-60	Sequence 60, Appl
63	5	1.8	14	9	US-10-024-123-8	Sequence 8, Appl
64	5	1.8	14	9	US-09-880-748-2438	Sequence 2438, Ap
65	5	1.8	14	10	US-09-205-658-290	Sequence 290, App
66	5	1.8	15	9	US-09-976-740-19	Sequence 19, Appl
67	5	1.8	15	9	US-10-157-669-34	Sequence 34, Appl
68	5	1.8	15	9	US-10-157-669-36	Sequence 36, Appl
69	5	1.8	15	10	US-09-962-055-19	Sequence 19, Appl
70	5	1.8	15	12	US-10-023-529-19	Sequence 19, Appl
71	5	1.8	15	12	US-10-023-523-19	Sequence 19, Appl
72	5	1.8	16	9	US-09-778-200-18	Sequence 18, Appl
73	5	1.8	16	9	US-09-778-200-21	Sequence 21, Appl
74	5	1.8	16	9	US-09-826-290-230	Sequence 230, App
75	5	1.8	16	9	US-09-826-290-341	Sequence 341, App
76	5	1.8	16	9	US-09-826-290-358	Sequence 358, App
77	5	1.8	16	9	US-09-976-740-25	Sequence 25, Appl
78	5	1.8	16	9	US-09-791-389-48	Sequence 48, Appl
79	5	1.8	16	10	US-09-791-393-48	Sequence 48, Appl
80	5	1.8	16	10	US-09-962-055-25	Sequence 25, Appl
81	5	1.8	16	10	US-09-791-378-101	Sequence 101, App
82	5	1.8	16	10	US-09-791-378-326	Sequence 326, App
83	5	1.8	16	12	US-10-023-529-25	Sequence 25, Appl
84	5	1.8	16	12	US-10-023-523-25	Sequence 25, Appl
85	5	1.8	17	9	US-09-805-337A-29	Sequence 29, Appl
86	5	1.8	18	10	US-09-864-761-40510	Sequence 40510, A
87	5	1.8	18	10	US-09-734-520-33	Sequence 33, Appl
88	5	1.8	18	12	US-10-012-034A-33	Sequence 33, Appl
89	5	1.8	29	10	US-09-168-474-19	Sequence 19, Appl
90	5	1.8	20	9	US-10-044-703-79	Sequence 79, Appl
91	5	1.8	20	10	US-09-864-761-38908	Sequence 38908, A
92	5	1.8	20	10	US-09-864-761-41108	Sequence 41108, A

```

93 5 1.8 20 10 US-09-864-761-47194
94 5 1.8 20 10 US-09-813-333-79
95 5 1.8 21 10 US-09-895-072-28
96 5 1.8 21 10 US-09-864-761-38136
97 5 1.8 22 10 US-09-864-761-38136
98 5 1.8 22 10 US-09-867-852-143
99 5 1.8 24 9 US-10-007-055-453
100 5 1.8 24 10 US-09-864-761-34620
101 5 1.8 24 10 US-09-864-761-47438
102 5 1.8 24 10 US-09-803-165-21
103 5 1.8 25 9 US-10-003-887-119
104 5 1.8 25 9 US-10-083-357-1227
105 5 1.8 25 9 US-10-007-280A-197
106 5 1.8 25 10 US-09-821-584-19
107 5 1.8 25 10 US-09-803-165-22
108 5 1.8 25 10 US-09-803-165-24
109 5 1.8 25 10 US-09-803-165-25
110 5 1.8 25 10 US-09-803-165-26
111 5 1.8 26 9 US-09-976-740-20
112 5 1.8 26 9 US-10-133-934-8
113 5 1.8 26 10 US-09-864-761-38833
114 5 1.8 26 10 US-09-562-055-20
115 5 1.8 26 12 US-10-023-529-20
116 5 1.8 26 12 US-10-023-529-20
117 5 1.8 27 10 US-09-864-761-34114
118 5 1.8 27 10 US-09-864-761-40933
119 5 1.8 28 9 US-09-941-548-16
120 5 1.8 28 9 US-09-976-740-26
121 5 1.8 28 10 US-09-864-761-40576
122 5 1.8 28 10 US-09-562-055-26
123 5 1.8 28 12 US-10-023-529-26
124 5 1.8 28 12 US-10-023-529-26
125 5 1.8 29 9 US-09-281-495-51
126 5 1.8 29 10 US-09-864-761-46130
127 5 1.8 30 9 US-09-969-730-177
128 5 1.8 30 10 US-09-864-761-33514
129 5 1.8 30 10 US-09-864-761-45776
130 5 1.8 30 10 US-09-732-091-16
131 5 1.8 31 9 US-09-774-039-136
132 5 1.8 31 10 US-09-050-516-44
133 5 1.8 31 10 US-09-764-887-159
134 5 1.8 31 10 US-09-864-761-41087
135 5 1.8 31 10 US-09-864-761-43875

```

ALIGNMENTS

```

RESULT 1
US-09-664-761-42347
: Sequence 42347, Application US/C9864761
: Patent No. US20020048763A
: GENERAL INFORMATION:
: APPLICANT: Penn. Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chan, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Acomica X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666

```

```

: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
: SEQ ID NO 42347
: LENGTH: 65
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC021881.2
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
: OTHER INFORMATION: EST_HUMAN HIT: AUI60252.1, EVALUJE 9.00e-24
: OTHER INFORMATION: SWISSPROT HIT: Q13206, EVALUJE 4.00e-24
US-09-864-761-42347

Query Match 2.8%; Score 8; DA 10; Length 65;
Best local Similarity 100.0%; Pred.No. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 DEEEEDAD 226
DB 45 DEEEEDAD 52

RESULT 2
US-09-826-290-48
: Sequence 48, Application US/09826290
: Patent No. US20020164668A1
: GENERAL INFORMATION:
: APPLICANT: Durham, L.Kathryn
: APPLICANT: Friedman, David L.
: APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri
: APPLICANT: Kimmel, Lida H.
: APPLICANT: Parekh, Rajesh Bhikhu
: APPLICANT: Potter, David M.
: APPLICANT: Rohlf, Christian
: APPLICANT: Silber, B. Michael
: APPLICANT: Stieger, Thomas R.
: APPLICANT: Sunderland, P. Irey
: APPLICANT: Townsend, Robert Reid
: APPLICANT: White, Frost
: APPLICANT: Williams, Stephen A.
: TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and

```

```

; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
; FILE REFERENCE: Alzheimer's Disease
; CURRENT APPLICATION NUMBER: US/69/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 62/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-26
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 13
; TYPE: PRT
; ORGANISM: homo sapien
US-09-826-290-48

Query Match      2.1%  Score 6; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  218 EDEEEE 223
Db  1 EDEEEE 6

RESULT 3
US-09-736-457-1855
; Sequence 1855, Application US/09/16457
; Patent No. US2002168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fast, Aijun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/03/736,457
; CURRENT FILING DATE: 2003-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1855
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-1855

Query Match      2.1%  Score 6; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  218 EDEEEE 223
Db  15 EDEEEE 20

RESULT 4
US-09-932-941-1855
; Sequence 1855, Application US/09/2941
; Patent No. US200212952A;
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watarabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.

```

```

; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andrea
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1855
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-1855

Query Match      2.1%  Score 6; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  218 EDEEEE 223
Db  15 EDEEEE 20

RESULT 5
US-09-849-626-1855
; Sequence 1855, Application US/09/849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clappett, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1855
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-1855

Query Match      2.1%  Score 6; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  218 EDEEEE 223
Db  15 EDEEEE 20

RESULT 6
US-10-017-754-1855
; Sequence 1855, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watarabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita

```

; APPLICANT: Carter, Darrick
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Bangor, Chaitanya S.
 ; APPLICANT: McNabb, Adria
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121-478C18
 ; CURRENT FILING DATE: 2001-10-29
 ; NUMBER OF SEQ ID NOS: 2004
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1855
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US 10-017-754-1855

Query Match 2.1%; Score 6; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 218 EDEEE 223

DB 15 EDEEE 20

RESULT 7
 US-10-057-505-25
 ; Sequence 25, Application US/10057505
 ; Patent No. US20020154674A1
 ; GENERAL INFORMATION:
 ; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
 ; APPLICANT: TSJEN, Roger
 ; APPLICANT: HEIM, Roger
 ; TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS
 ; FILE REFERENCE: REG1260-3
 ; CURRENT APPLICATION NUMBER: US/10/057-505
 ; CURRENT FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: US 08/792,553
 ; PRIOR FILING DATE: 1997-01-31
 ; PRIOR APPLICATION NUMBER: US 09/396,003
 ; PRIOR FILING DATE: 1999-09-13
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 25
 ; LENGTH: 25
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Linker moiety
 ; US-10-057-505-25

Query Match 2.1%; Score 6; DB 9; Length 25;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 209 DLYDDD 214

DB 12 DLYDDD 17

RESULT 8
 US-09-866-538-21
 ; Sequence 21, Application US/09866538
 ; Publication No. US2003032088A1
 ; GENERAL INFORMATION:
 ; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
 ; APPLICANT: TSJEN, Roger
 ; APPLICANT: Campbell, Robert
 ; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS

; FILE REFERENCE: REG1530-2
 ; CURRENT APPLICATION NUMBER: US/09/866,538
 ; CURRENT FILING DATE: 2001-05-24
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 21
 ; LENGTH: 33
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: ADDITIONAL 33 AMINO ACIDS TAG TO THE N-TERMINUS OF THE GPPS
 ; US-09-866-538-21

Query Match 2.1%; Score 6; DB 9; Length 33;
 Best Local Similarity 100.0%; Pred. No. 11e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 209 DLYDDD 214

DB 24 DLYDDD 29

RESULT 9
 US-10-121-258-22
 ; Sequence 22, Application US/10121258
 ; Publication No. US20030059835A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Campbell, Robert
 ; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
 ; FILE REFERENCE: UC083.1CP2CP1
 ; CURRENT APPLICATION NUMBER: US/10/121-258
 ; CURRENT FILING DATE: 2002-04-10
 ; PRIOR APPLICATION NUMBER: 99/794,308
 ; PRIOR FILING DATE: 2001-02-26
 ; PRIOR APPLICATION NUMBER: 09/866,538
 ; PRIOR FILING DATE: 2001-05-24
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 22
 ; LENGTH: 33
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: 6xHis Tag
 ; US-10-121-258-22

Query Match 2.1%; Score 6; DB 9; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 209 DLYDDD 214

DB 24 DLYDDD 29

RESULT 10
 US-09-864-761-46817
 ; Sequence 46817, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Acomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GH 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 4917
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 46817
 ; LENGTH: 34
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AP000076.1
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 0.91
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 1.5
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 1.5

Query Match 2.18; Score 6; DB 10; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1.2e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 IKFPID 67
 |||||
 Db 6 IKFPID 11

RESULT 11
 US-09-864-761-48968
 ; Sequence 48968, Application US/09864761
 ; Patent No. US20020048763A2
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Accmca-X 1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GH 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 4917
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 48968
 ; LENGTH: 34
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AP000076.1
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL - 1.8
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL - 1.8
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL - 1.8

Query Match 2.18; Score 6; DB 10; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1.2e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 IKFPID 67
 |||||
 Db 6 IKFPID 11

RESULT 12
 US-10-083-815-68
 ; Sequence 68, Application US/10083815
 ; Publication No. US20030026781A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, Christen M.
 ; APPLICANT: Cleveland, William
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING
 ; TITLE OF INVENTION: ENDOGENOUS INHIBITOR OF ATP SYNTHASE, INCLUDING
 ; TITLE OF INVENTION: TREATMENT FOR DIABETES
 ; FILE REFERENCE: 660088.435C2
 ; CURRENT APPLICATION NUMBER: US/10/083,815
 ; CURRENT FILING DATE: 2002-02-27
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 68
 ; LENGTH: 35
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence


```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/884,687
FILING DATE: 19-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/679,865
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John S.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 02307Z-069000
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-884-681-39

Query Match      2.1%; Score 6; DB 10; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 DLYDDD 214
DB 24 DLYDDD 29

RESULT 16
US-08-424-550B-588
: Sequence 588, Application US/08424550B
: Patent No. US20020119447A1
: GENERAL INFORMATION:
: APPLICANT: JOHN N. SIMONS
: APPLICANT: TAMI J. PILOT-MATIAS
: APPLICANT: GEORGE J. DAWSON
: APPLICANT: GEORGE G. SCHLAUDER
: APPLICANT: SURESH M. DESAI
: APPLICANT: THOMAS P. LEARY
: APPLICANT: ANTHONY SCOTT MUEHRHOFF
: APPLICANT: JAMES G. ERKER
: APPLICANT: SHERI L. BULCK
: APPLICANT: ISA K. MUSHAMMAR
: TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
: NUMBER OF SEQUENCES: 716
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ABBOTT LABORATORIES D377/APSD
: STREET: 100 ABBOTT PARK ROAD
: CITY: ABBOTT PARK
: STATE: IL
: COUNTRY: USA
: ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,550B
FILING DATE:
CLASSIFICATION: 435435
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E.

```

```

: REGISTRATION NUMBER: 33,207
: REFERENCE/DOCKET NUMBER: 5527-PC-01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 708-937-6365
: TELEFAX: 708-938-2623
: INFORMATION FOR SEQ ID NO: 588:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 45 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-424-550H-588

Query Match      2.1%; Score 6; DB 8; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GKGASG 260
DB 36 GKGASG 41

RESULT 17
US-09-864-761-36758
: Sequence 36758, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Rank, David K.
: APPLICANT: Penn, Sharron G.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
: FILE REFERENCE: Aecomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117

```

```

? SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 36758
? LENGTH: 47
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC009653.2
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 1.7
? OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 2.6
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 1
? OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL - 1.7
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.1
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 0.9
? OTHER INFORMATION: SWISSPROT HIT: P54105, EVALU6 5.00e-03
? OTHER INFORMATION: EST_HUMAN HIT: A0123910.1, EVALU6 6.90e-02
US-09-864-761-39559

```

```

Query Match          2.1% Score 6; DB 10; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.5e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 219 DEEDED 224
    |||
DB 14 DEEDED 19

```

RESULT 18

```

US-09-864-761-39559
? Sequence 39559, Application US/09864761
? Patent No. US20020048763A1
? GENERAL INFORMATION:
? APPLICANT: Penn, Sharron G.
? APPLICANT: Rank, David R.
? APPLICANT: Hanzel, David K.
? APPLICANT: Chen, Wensheng
? TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
? FILE REFERENCE: Acomica-X-1
? CURRENT APPLICATION NUMBER: US/09/864,761
? PRIOR FILING DATE: 2001-05-23
? PRIOR APPLICATION NUMBER: US 60/180,312
? PRIOR FILING DATE: 2000-02-04
? PRIOR APPLICATION NUMBER: US 60/207,456
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: US 09/632,366
? PRIOR FILING DATE: 2000-08-03
? PRIOR APPLICATION NUMBER: US 60/236,359
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: PCT/US01/00666
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/608,408

```

```

? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 39559
? LENGTH: 49
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC003685.1
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1.8
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 2.5
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 1.5
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.4
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 1.4
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 1.3
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 3.9
US-09-864-761-39559

```

```

Query Match          2.1% Score 6; DB 10; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.6e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 219 DEEDED 224
    |||||
DB 10 DEEDED 15

```

RESULT 19

```

US-09-864-761-33667
? Sequence 33667, Application US/09864761
? Patent No. US20020048763A1
? GENERAL INFORMATION:
? APPLICANT: Penn, Sharron G.
? APPLICANT: Rank, David R.
? APPLICANT: Hanzel, David K.
? APPLICANT: Chen, Wensheng
? TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
? FILE REFERENCE: Acomica-X-1
? CURRENT APPLICATION NUMBER: US/09/864,761
? PRIOR FILING DATE: 2001-05-23
? PRIOR APPLICATION NUMBER: US 60/180,312
? PRIOR FILING DATE: 2000-02-04
? PRIOR APPLICATION NUMBER: US 60/207,456
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: US 09/632,366
? PRIOR FILING DATE: 2000-08-03
? PRIOR APPLICATION NUMBER: GB 24263.6
? PRIOR FILING DATE: 2000-10-04
? PRIOR APPLICATION NUMBER: US 60/236,359
? PRIOR FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: PCT/US01/00666
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33667
; LENGTH: 52
; TYPE: PRI
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO APC0007.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL - 5.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 2
; OTHER INFORMATION: EXPRESSED IN HEP100, SIGNAL - 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 1.2
; OTHER INFORMATION: EXPRESSED IN HEPA1, SIGNAL - 1.8
; OTHER INFORMATION: EST HUMAN HIT: AW779776.1, EVALUATE 1.00e-07
; OTHER INFORMATION: SW:SPROT HIT: Q05086, EVALUATE 2.00e-09
; OTHER INFORMATION: EST HUMAN HIT: BE564683.1, EVALUATE 3.00e-08
US-09-864-761-33667

Query Match      2.1%  Score 6;  DB 10; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.7e-02;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  16 ELKSEQ 21
DB   31 ELKSEQ 16

RESULT 20
US-09-864-761-39215
; Sequence 39215, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David K.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US 09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39215
; LENGTH: 54
; TYPE: PRI
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003049.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 5.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 5.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 6.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 5.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 5.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 5.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 5.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 5.5
US-09-864-761-39215

Query Match      2.1%  Score 6;  DB 10; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.8e-02;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  218 EDEEEE 223
DB   12 EDEEEE 17

RESULT 21
US-09-864-761-44300
; Sequence 44300, Application: US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David K.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 43117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44300
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004123.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 0.83
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 0.95
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 0.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 0.85
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 0.87
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 0.87
US-09-864-761-44300

Query Match          2.1%      Score 6; DB 10; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0
QY 218 EDEEE 223
DB 12 EDEEE 17
|||||
|||||

RESULT 22
US-09-864-761-44147
; Sequence 44147, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David K.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,353
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44147
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005686.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 0.66
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 0.49
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 0.45
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 0.45
; OTHER INFORMATION: SWISSPROT HIT: 000233, EVALUAE 4.60e+00
US-09-864-761-44147

Query Match          2.1%      Score 6; DB 10; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0
QY 122 TTTLSV 127
DB 39 TTTLSV 44
|||||
|||||

RESULT 23
US-09-864-761-44799
; Sequence 44799, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David K.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,353
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30

```

```

; PRICK APPLICATION NUMBER: PCT/US01/00664
; PRICK FILING DATE: 2001-01-30
; PRICK APPLICATION NUMBER: PCT/US01/00669
; PRICK FILING DATE: 2001-01-30
; PRICK APPLICATION NUMBER: PCT/US01/00665
; PRICK FILING DATE: 2001-01-30
; PRICK APPLICATION NUMBER: PCT/US01/00668
; PRICK FILING DATE: 2001-01-30
; PRICK APPLICATION NUMBER: PCT/US01/00663
; PRICK FILING DATE: 2001-01-30
; PRICK APPLICATION NUMBER: PCT/US01/00662
; PRICK FILING DATE: 2001-01-30
; PRICK APPLICATION NUMBER: PCT/US01/00661
; PRICK FILING DATE: 2001-01-30
; PRICK APPLICATION NUMBER: PCT/US01/00670
; PRICK FILING DATE: 2001-01-30
; PRICK APPLICATION NUMBER: US 60/234,587
; PRICK FILING DATE: 2000-09-21
; PRICK APPLICATION NUMBER: US 09/608,408
; PRICK FILING DATE: 2000-06-30
; PRICK APPLICATION NUMBER: US 09/774,203
; PRICK FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44795
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL158088.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 0.86
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 0.82
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 0.76
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 0.76
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 0.64
; US-09-964 751-44795

```

```

Query Match 2.1% Score 6; Dh 10; Length 57;
Best local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

```

```

QY 218 EDEEE 223
DB 13 EDEEE 18

```

```

RESULT 24
US-09 718-626-3524
; Sequence 3524, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIROGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/155162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0

```

```

; SEQ ID NO 3524
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-3524

```

```

Query Match 2.1% Score 6; Dh 9; Length 60;
Best local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

```

```

QY 8 SSOKAL 13
DB 6 SSOKAL 11

```

```

RESULT 25
US-09-864-761-47146
; Sequence 47146, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Rank, Sharron G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL IN
; FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,455
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.5
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,587
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47146
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004806.1

```

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 0.62
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 0.61
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 0.68
 OTHER INFORMATION: SWISSPROT HIT: P45257, EVALUATE 1.20e-00
 OTHER INFORMATION: EST_HUMAN HIT: W00669.1, EVALUATE 2.00e-08
 US-09-864-761-47146

Query Match 2.1% Score 6; DB 10; Length 62;
 Best Local Similarity 100.0%; Pred. No. 2e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 200 DNASISL 205
 DB 16 DNASISL 21

RESULT 26

US-09-932-679-56
 Sequence 56, Application: US/09332679
 Patent No. US20020358801A1
 GENERAL INFORMATION:
 APPLICANT: Lok, Si
 APPLICANT: Conklin, Barrell C.
 APPLICANT: No. US20020358801A1ak, Julia B.
 TITLE OF INVENTION: Mammalian Alpha Helical Protein-1
 FILE REFERENCE: 97-7101
 CURRENT APPLICATION NUMBER: US/09/932,679
 PRIOR FILING DATE: 2001-08-16
 PRIOR APPLICATION NUMBER: 09/209,525
 PRIOR FILING DATE: 1998-12-10
 PRIOR APPLICATION NUMBER: 60/067,779
 PRIOR FILING DATE: 1997-12-10
 NUMBER OF SEQ ID NOS: 56
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 56
 TYPE: PRT
 LENGTH: 65
 ORGANISM: Homo sapiens
 US-09-932-679 56

Query Match 2.1% Score 6; DB 10; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.1e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 219 DEEED 224
 DB 58 DEEED 63

RESULT 27

US-09-912 962-35
 Sequence 35, Application US/09912962
 Patent No. US2002007619A1
 GENERAL INFORMATION:
 APPLICANT: de Lange, Iltia
 Broccoli, Dominique
 Smogorzewska, Agata
 TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND
 DIAGNOSTIC AND THERAPEUTIC USE THEREOF
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESS: KLAUHER & JACKSON
 STREET: 411 Backensack Avenue
 CITY: Backensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/912,962
 FILING DATE: 25-Jul-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/018,635
 FILING DATE: 04-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: David A. Jackson
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-142 CIP1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-487-5800
 TELEFAX: 201-343-1884
 TELEX: 133521

INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 67 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 SEQUENCE DESCRIPTION: SEQ ID NO: 33:
 US-09-912-962-33

Query Match 2.1% Score 6; DB 10; Length 67;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 220 EEEEDA 225
 DB 59 EEEEDA 64

RESULT 28

US-09-738-626-5114
 Sequence 5114, Application US/09738626
 Publication No. US20020197605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SAIOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO 5114
 LENGTH: 70
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-5114

Query Match 2.1% Score 6; DB 9; Length 70;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 221 EECDA 226

|||||

Db 11 EECDA 16

RESULT 29

US-09-727-963A-20

; Sequence 20, Application US/0977963A

; Patent No. US2002015106A1

; GENERAL INFORMATION:

; APPLICANT: V.I. Technologies, Inc.

; APPLICANT: Hammond, David J.

; TITLE OF INVENTION: METHOD OF IDENTIFYING A LIGAND FOR A TARGET MOLECULE

; FILE REFERENCE: 18242-505

; CURRENT APPLICATION NUMBER: US/09/727,963A

; PRIOR FILING DATE: 2000-12-01

; PRIOR APPLICATION NUMBER: U.S.S.N. 60/327,557

; NUMBER OF SEQ ID NOS: 99

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 20

; LENGTH: 6

; TYPE: PRI

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence : peptide

; OTHER INFORMATION: ligand

US-09-727-963A-20

Query Match

Best Local Similarity 100.0%; DB 9; Length 6;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LLSVI 128

|||||

Db 2 LLSVI 6

RESULT 30

US-09-910-346C-10

; Sequence 10, Application US/09910346C

; Publication No. US20030027752A1

; GENERAL INFORMATION:

; APPLICANT: STEWARD, LANCE E

; APPLICANT: FERNANDEZ-SALAS, ESTER

; APPLICANT: HERRINGTON, TODD M

; APPLICANT: AOKI, KEI R

; TITLE OF INVENTION: Leucine-based motif and clostridial neurotoxins

; FILE REFERENCE: D-2885C1P

; CURRENT APPLICATION NUMBER: US/09/910,346C

; PRIOR FILING DATE: 2000-07-21

; PRIOR APPLICATION NUMBER: US 09/520,840

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 10

; LENGTH: 7

; TYPE: PRI

; ORGANISM: rat

US-09-910-346C-10

Query Match

Best Local Similarity 100.0%; DB 9; Length 7;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 KVVLL 242

|||||

Db 3 KVVLL 7

RESULT 31

US-10-083-815-2

; Sequence 2, Application US/0083815

; Publication No. US20030026781A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, Christen M.

; APPLICANT: Cleveland, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING

; TITLE OF INVENTION: ENDOGENOUS INHIBITOR OF ATP SYNTHASE, INCLUDING

; TITLE OF INVENTION: TREATMENT FOR DIABETES

; FILE REFERENCE: 660088.435C2

; CURRENT APPLICATION NUMBER: US/10/083,815

; CURRENT FILING DATE: 2002-02-27

; NUMBER OF SEQ ID NOS: 72

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 7

; TYPE: PRI

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Epitope tag

US-10-083-815-2

Query Match

Best Local Similarity 100.0%; DB 9; Length 7;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 YDDDD 232

|||||

Db 2 YDDDD 6

RESULT 32

US-09-229-751A-40

; Sequence 40, Application US/09229751A

; Publication No. US20030044838A1

; GENERAL INFORMATION:

; APPLICANT: Turnbough, Charles K

; TITLE OF INVENTION: PEPTIDE LIGANDS THAT BIND TO SURFACES

; TITLE OF INVENTION: OF BACTERIAL CELLS

; NUMBER OF SEQUENCES: 80

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Glenna Hendricks

; STREET: P.O. Box 2509

; CITY: Fairfax

; STATE: VA

; COUNTRY: USA

; ZIP: 22031

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/229,751A

; FILING DATE: 14-Jan-1999

; CLASSIFICATION: <unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Hendricks, Glenna M

; REGISTRATION NUMBER: 32,535

; REFERENCE/DOCKET NUMBER: turn

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 425-8405

; TELEFAX: (703) 425-8406

; INFORMATION FOR SEQ ID NO: 40:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; SEQUENCE DESCRIPTION: SEQ ID NO: 40:

US-09-229-751A-40


```
Query Match: 1.8%, Score 5; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 QNVR: 122
DB 3 QNVR: 7

RESULT 33
US-09-989-789-733
; Sequence 733, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 733
; LENGTH: 7
; TYPE: PRI
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-733

Query Match: 1.8%, Score 5; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 NSSDL 205
DB 1 NSSDL 5

RESULT 34
US-09-989-789-1017
; Sequence 1017, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1017
; LENGTH: 7
; TYPE: PRI
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-1017

Query Match: 1.8%, Score 5; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 QSGEL 109
DB 1 QSGEL 5

RESULT 35
US-09-989-789-1222
; Sequence 1222, Application US/09989789
; Patent No. US20020063379A1
```

```
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1222
; LENGTH: 7
; TYPE: PRI
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-1222

Query Match: 1.8%, Score 5; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 QSGEL 109
DB 1 QSGEL 5

RESULT 36
US-09-989-789-1231
; Sequence 1231, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1231
; LENGTH: 7
; TYPE: PRI
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP
US-09-989-789-1231

Query Match: 1.8%, Score 5; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 QSGEL 109
DB 1 QSGEL 5

RESULT 37
US-10-010-184A-8
; Sequence 8, Application US/10010184A
; Publication No. US20030008828A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Pharma Company
; APPLICANT: Priestly, et al.
; TITLE OF INVENTION: No. US20030008828A1c1 Lactam Inhibitors of Hepatitis C Virus
; FILE REFERENCE: PH-7087-A
; CURRENT APPLICATION NUMBER: US/10/010,184A
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 09/626,286
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 8
; LENGTH: 9
```

```

: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: The synthesis of this peptide may be performed on an ABI 43A peptide synthesizer using readily available materials well known to
: OTHER INFORMATION: routinely skilled artisans
: FEATURE:
: NAME/KEY: ACETYLATION
: LOCATION: (1)-(2)
: OTHER INFORMATION: acetyl group
: FEATURE:
: NAME/KEY: MOD_RES
: LOCATION: (3)-(3)
: OTHER INFORMATION: Aspartic acid modified with EDANS, 5-[(2'-aminoethyl)amino]naphthalene-1-sulfonate
: FEATURE:
: NAME/KEY: MOD_RES
: LOCATION: (5)-(5)
: OTHER INFORMATION: Lysine modified by Dabcyl, 4-[[4'-(dimethylamino)phenyl]azo]benzyl
: OTHER INFORMATION: c acid
: US-10-010-184A-8

```

```

Query Match      1.8%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 217 DEEED 221
DB 1 DEEED 5

```

```

RESULT 38
US-10-102-283-79
: Sequence 79, Application US/10102283
: Publication No. US20030027181A1
: GENERAL INFORMATION:
: APPLICANT: O'Brien, Timothy J.
: APPLICANT: Cannon, Martin J.
: APPLICANT: Santin, Alessandro
: TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
: FILE REFERENCE: D6223CIP/A/O/CIP2
: CURRENT APPLICATION NUMBER: US/10/102,283
: CURRENT FILING DATE: 2002-03-20
: PRIOR FILING DATE: 2001-07-30
: NUMBER OF SEQ ID NOS: 190
: SEQ ID NO 79
: LENGTH: 9
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Residues 293-301 of the hepsin protein
: US-10-102-283-79

```

```

Query Match      1.8%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 274 GWGNT 278
DB 3 GWGNT 7

```

```

RESULT 39
US-10-102-283-176
: Sequence 176, Application US/10102283
: Publication No. US20030027181A1
: GENERAL INFORMATION:
: APPLICANT: O'Brien, Timothy J.

```

```

: APPLICANT: Cannon, Martin J.
: APPLICANT: Santin, Alessandro
: TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
: FILE REFERENCE: D6223CIP/A/O/CIP2
: CURRENT APPLICATION NUMBER: US/10/102,283
: CURRENT FILING DATE: 2002-03-20
: PRIOR FILING DATE: 2001-07-30
: NUMBER OF SEQ ID NOS: 190
: SEQ ID NO 176
: LENGTH: 9
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Residues 294-302 of the hepsin protein
: US-10-102-283-176

```

```

Query Match      1.8%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 274 GWGNT 278
DB 2 GWGNT 6

```

```

RESULT 40
US-10-102-283-182
: Sequence 182, Application US/10102283
: Publication No. US20030027181A1
: GENERAL INFORMATION:
: APPLICANT: O'Brien, Timothy J.
: APPLICANT: Cannon, Martin J.
: APPLICANT: Santin, Alessandro
: TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
: FILE REFERENCE: D6223CIP/A/O/CIP2
: CURRENT APPLICATION NUMBER: US/10/102,283
: CURRENT FILING DATE: 2002-03-20
: PRIOR FILING DATE: 2001-07-30
: NUMBER OF SEQ ID NOS: 190
: SEQ ID NO 182
: LENGTH: 9
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Residues 293-301 of the hepsin protein
: US-10-102-283-182

```

```

Query Match      1.8%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 274 GWGNT 278
DB 3 GWGNT 7

```

```

Search completed: April 10, 2003, 10:43:39
Job time : 14.28 secs

```



```

101 5 1.8 9 4 US-09-510-738A-79 Sequence 79, Appl
102 5 1.8 9 4 US-09-510-738A-176 Sequence 176, App
103 5 1.8 9 4 US-09-510-738A-182 Sequence 182, App
104 5 1.8 9 4 US-09-586-472-74 Sequence 74, Appl
105 5 1.8 9 4 US-09-528-706-74 Sequence 74, Appl
106 5 1.8 10 1 US-08-199-776-8 Sequence 8, Appl
107 5 1.8 10 1 US-08-199-776-9 Sequence 9, Appl
108 5 1.8 10 1 US-08-199-776-10 Sequence 10, Appl
109 5 1.8 10 1 US-08-199-776-11 Sequence 11, Appl
110 5 1.8 10 1 US-08-199-776-12 Sequence 12, Appl
111 5 1.8 10 1 US-08-199-776-13 Sequence 13, Appl
112 5 1.8 10 1 US-08-199-776-14 Sequence 14, Appl
113 5 1.8 10 1 US-08-199-776-17 Sequence 17, Appl
114 5 1.8 10 2 US-08-337-646A-30 Sequence 30, Appl
115 5 1.8 10 2 US-08-574-959A-17 Sequence 17, Appl
116 5 1.8 10 3 US-09-139-762A-48 Sequence 48, Appl
117 5 1.8 10 3 US-09-139-762A-87 Sequence 87, Appl
118 5 1.8 10 3 US-09-139-762A-116 Sequence 116, App
119 5 1.8 10 3 US-08-663-731-8 Sequence 8, Appl
120 5 1.8 10 3 US-08-663-731-9 Sequence 9, Appl
121 5 1.8 10 3 US-08-663-731-10 Sequence 10, Appl
122 5 1.8 10 3 US-08-663-731-11 Sequence 11, Appl
123 5 1.8 10 3 US-08-663-731-12 Sequence 12, Appl
124 5 1.8 10 3 US-08-663-731-13 Sequence 13, Appl
125 5 1.8 10 3 US-08-663-731-14 Sequence 14, Appl
126 5 1.8 10 3 US-08-663-731-17 Sequence 17, Appl
127 5 1.8 10 3 US-08-875-338-8 Sequence 8, Appl
128 5 1.8 10 3 US-08-875-338-9 Sequence 9, Appl
129 5 1.8 10 3 US-08-875-338-10 Sequence 10, Appl
130 5 1.8 10 3 US-08-875-338-11 Sequence 11, Appl
131 5 1.8 10 3 US-08-875-338-12 Sequence 12, Appl
132 5 1.8 10 3 US-08-875-338-13 Sequence 13, Appl
133 5 1.8 10 3 US-08-875-338-14 Sequence 14, Appl
134 5 1.8 10 3 US-08-875-338-17 Sequence 17, Appl
135 5 1.8 10 3 US-09-020-116-3 Sequence 3, Appl

```

ALIGNMENTS

```

RESULT 1
US-08-080-073 35
; Sequence 35, Application US/0808073
; Patent No. 5484255
; GENERAL INFORMATION:
; APPLICANT: Clechanover, Aaron J.
; APPLICANT: Blumenfeld, Nava
; APPLICANT: Gonen, Hedva
; TITLE OF INVENTION: Ubiquitin Carrier Enzyme E2-F1,
; TITLE OF INVENTION: Purification, Production, and Use
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release 1.0, Version 1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/080,073
; FILING DATE: 21-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1448.005000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600

```

```

; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-080-073-35

```

```

Query Match 4.38: Score 12; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 124 LLSVISLLNEPN 135
|||||
Db 5 LLSVISLLNEPN 16
|||||

```

```

RESULT 2
US-08-318-837-24
; Sequence 24, Application US/083:8837
; Patent No. 5981277
; GENERAL INFORMATION:
; APPLICANT: FRANSEN, LUCIA; DEVOS, KATHLEEN; VAN DE VOORDE,
; APPLICANT: ANDRE; VAN HEUVERSWIJN, HUGO
; TITLE OF INVENTION: NEW POLYPEPTIDES AND PEPTIDES, NUCLEIC ACID
; TITLE OF INVENTION: CODING FOR THEM, AND THEIR USE IN THE FIELD OF TUMOR THERAP
; TITLE OF INVENTION: IMMUNOLOGY
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIEMAN AND MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,837
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP 93/01022
; FILING DATE: 28-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 92 401.231.3
; FILING DATE: 30-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410.007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Mouse, human
; CELL LINE: PUS-1.8, THP-1
US-08-318-837-24

```

```

Query Match 2.58: Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 134 PNIESPA 240
 111 1111
 DB 16 PNIESPA 22

RESULT 3

US-08-499-676A-31
 : Sequence 31, Application US/08499676A
 : Patent No. 5998154
 : GENERAL INFORMATION:
 : APPLICANT: AGNES SCHONBRUNN
 : TITLE OF INVENTION: SOMATOSTATIN RECEPTOR PEPTIDE
 : NUMBER OF SEQUENCES: 43
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
 : STREET: 1177 West Loop South, 10th Floor
 : CITY: Houston
 : STATE: TX
 : COUNTRY: USA
 : ZIP: 77027-9095
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/499,676A
 : FILING DATE: July 7, 1995
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER:
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Krieger, Paul H.
 : REGISTRATION NUMBER: 25,866
 : REFERENCE/DOCKET NUMBER: 79247.3/A9517505
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 713-858-0829
 : TELEFAX: 713-850-0165
 : INFORMATION FOR SEQ ID NO: 31:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 43 amino acids
 : TYPE: amino acid
 : STRANDEDNESS:
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : US-08-499-676A-31

Query Match 2.5% Score 7; DB 2; Length 43;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 224
 11 1111
 DB 12 EDEEEE 18

RESULT 4

US-08 816-346-10
 : Sequence 10, Application US/08816346
 : Patent No. 6127525
 : GENERAL INFORMATION:
 : APPLICANT: Crystal, Ronald G.
 : APPLICANT: Falck-Pedersen, Erik
 : APPLICANT: Gall, Jason
 : APPLICANT: Kovesdi, Imre
 : APPLICANT: Wickham, Thomas J.
 : TITLE OF INVENTION: CHIMERIC ADENOVIRAL COAT PROTEIN AND
 : NUMBER OF SEQUENCES: 60
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.

: STREET: TWO PRUDENTIAL PLAZA - 4900
 : CITY: CHICAGO
 : STATE: ILLINOIS
 : COUNTRY: USA
 : ZIP: 60601-6780
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/816,346
 : FILING DATE:
 : CLASSIFICATION: 530
 : ATTORNEY/AGENT INFORMATION:
 : REFERENCE/DOCKET NUMBER: 67167
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312/616-5600
 : TELEFAX: 312/616-5700
 : TELEX: 25-3533
 : INFORMATION FOR SEQ ID NO: 10:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 51 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : US-08-816-346-10

Query Match 2.5% Score 7; DB 3; Length 51;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 224
 11 1111
 DB 10 EDEEEE 16

RESULT 5

US-09-335-411-10
 : Sequence 10, Application US/09335411
 : Patent No. 6153435
 : GENERAL INFORMATION:
 : APPLICANT: Crystal, Ronald G.
 : APPLICANT: Falck-Pedersen, Erik
 : APPLICANT: Gall, Jason
 : APPLICANT: Kovesdi, Imre
 : APPLICANT: Wickham, Thomas J.
 : TITLE OF INVENTION: CHIMERIC ADENOVIRAL COAT PROTEIN AND
 : NUMBER OF SEQUENCES: 60
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
 : STREET: TWO PRUDENTIAL PLAZA - 4900
 : CITY: CHICAGO
 : STATE: ILLINOIS
 : COUNTRY: USA
 : ZIP: 60601-6780
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/335,411
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/816,346
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : REFERENCE/DOCKET NUMBER: 67167
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312/616-5600

TELEFAX: 312/616-6700
TELEX: 25-3533
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-335-411-10

Query Match 2.18; Score 7; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 218 EDEEE 224
IIIIII
DB 10 EDEEE 16

RESULT 6
US-08-199-776-15
Sequence 15, Application US/08199776
Patent No. 5594120
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 5594120el integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/199,776
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: H0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens

Query Match 2.18; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 218 EDEEE 223
I..III
DB 5 EDEEE 10

RESULT 7
US-08-199-776-16
Sequence 16, Application US/08199776
Patent No. 5594120
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 5594120el integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/199,776
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: H0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens

Query Match 2.18; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 218 EDEEE 223
IIIIII
DB 4 EDEEE 9

RESULT 8
US-08-663-731-15
Sequence 15, Application US/08663731
Patent No. 6057423
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 6057423el integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:

```
1 MEDIUM TYPE: Floppy disk
2 COMPUTER: IBM PC compatible
3 OPERATING SYSTEM: PC-DOS/MS-DOS
4 SOFTWARE: Patent Release #1.0, Version #1.25
5 CURRENT APPLICATION DATA: US/08/663,731
6 FILING DATE:
7 CLASSIFICATION:
8 PRIOR APPLICATION DATA:
9 APPLICATION NUMBER: 08/199,776
10 FILING DATE:
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Plumer, Elizabeth R.
13 REGISTRATION NUMBER: 36,637
14 REFERENCE/DOCKET NUMBER: B0801/7020
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 617-720-3500
17 TELEFAX: 617-720-2441
18 INFORMATION FOR SEQ ID NO: 15:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH: 10 amino acids
21 TYPE: amino acid
22 STRANDEDNESS: single
23 TOPOLOGY: linear
24 MOLECULE TYPE: peptide
25 HYPOTHETICAL: YES
26 ANTI-SENSE: NO
27 FRAGMENT TYPE: internal
28 ORIGINAL SOURCE:
29 ORGANISM: Homo sapiens
30 US-08-663-731-15
31
32 Query Match 2.1% Score 6; DB 3; Length 10;
33 Best Local Similarity 100.0%; Pred. No. 25;
34 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
35
36 QY 218 EDEEEE 223
37 Db 5 EDEEEE 10
38
39 RESULT 9
40 US-08-663-731-16
41 Sequence 16, Application US/08663731
42 Patent No. 6657423
43 GENERAL INFORMATION:
44 APPLICANT: Brenner, Michael B.
45 APPLICANT: Parker, Christina M.
46 TITLE OF INVENTION: No. 6057423el Integrin alpha subunit
47 NUMBER OF SEQUENCES: 25
48 CORRESPONDENCE ADDRESS:
49 ADDRESSEE: Woll, Greenfield and Sacks, P.C.
50 STREET: 600 Atlantic Avenue
51 CITY: Boston
52 STATE: MA
53 COUNTRY: USA
54 ZIP: 02210
55 COMPUTER READABLE FORM:
56 MEDIUM TYPE: Floppy disk
57 COMPUTER: IBM PC compatible
58 OPERATING SYSTEM: PC-DOS/MS-DOS
59 SOFTWARE: Patent Release #1.0, Version #1.25
60 CURRENT APPLICATION DATA:
61 APPLICATION NUMBER: US/08/663,731
62 FILING DATE:
63 CLASSIFICATION:
64 PRIOR APPLICATION DATA:
65 APPLICATION NUMBER: 08/199,776
66 FILING DATE:
67 ATTORNEY/AGENT INFORMATION:
68 NAME: Plumer, Elizabeth R.
69 REGISTRATION NUMBER: 36,637
70 REFERENCE/DOCKET NUMBER: B0801/7020
```

```
1 TELECOMMUNICATION INFORMATION:
2 TELEPHONE: 617-720-3500
3 TELEFAX: 617-720-2441
4 INFORMATION FOR SEQ ID NO: 16:
5 SEQUENCE CHARACTERISTICS:
6 LENGTH: 10 amino acids
7 TYPE: amino acid
8 STRANDEDNESS: single
9 TOPOLOGY: linear
10 MOLECULE TYPE: peptide
11 HYPOTHETICAL: YES
12 ANTI-SENSE: NO
13 FRAGMENT TYPE: internal
14 ORIGINAL SOURCE:
15 ORGANISM: Homo sapiens
16 US-08-663-731-16
17
18 Query Match 2.1% Score 6; DB 3; Length 10;
19 Best Local Similarity 100.0%; Pred. No. 25;
20 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
21
22 QY 218 EDEEEE 223
23 Db 4 EDEEEE 9
24
25 RESULT 10
26 US-08-879-338-15
27 Sequence 15, Application US/08879338A
28 Patent No. 6063906
29 GENERAL INFORMATION:
30 APPLICANT: Brenner, Michael B.
31 APPLICANT: Parker, Christina M.
32 TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
33 TITLE OF INVENTION: Subunit
34 FILE REFERENCE: B0801/7080/ERP
35 CURRENT APPLICATION NUMBER: US/08/879,338A
36 CURRENT FILING DATE: 1997-06-20
37 EARLIER APPLICATION NUMBER: US 08/663,731
38 EARLIER FILING DATE: 1996-06-14
39 EARLIER APPLICATION NUMBER: US 08/199,776
40 EARLIER FILING DATE: 1994-02-18
41 NUMBER OF SEQ ID NOS: 31
42 SOFTWARE: FastSeq for Windows Version 3.0
43 SEQ ID NO 15
44 LENGTH: 10
45 TYPE: PRT
46 ORGANISM: Homo Sapiens
47 US-08-879-338-15
48
49 Query Match 2.1% Score 6; DB 3; Length 10;
50 Best Local Similarity 100.0%; Pred. No. 25;
51 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
52
53 QY 218 EDEEEE 223
54 Db 5 EDEEEE 10
55
56 RESULT 11
57 US-08-879-338-16
58 Sequence 16, Application US/08879338A
59 Patent No. 6063906
60 GENERAL INFORMATION:
61 APPLICANT: Brenner, Michael B.
62 APPLICANT: Parker, Christina M.
63 TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
64 TITLE OF INVENTION: Subunit
65 FILE REFERENCE: B0801/7080/ERP
66 CURRENT APPLICATION NUMBER: US/08/879,338A
67 CURRENT FILING DATE: 1997-06-20
68 EARLIER APPLICATION NUMBER: US 08/663,731
69 EARLIER FILING DATE: 1996-06-14
```

EARLIER APPLICATION NUMBER: US 08/199,776
EARLIER FILING DATE: 1994-02-18
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 16
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapiens
US-08-879-338-16

Query Match 2.1% Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 218 EDEEEE 223
IIIII
DB 4 EDEEEE 9

RESULT 12

US-09-293-238B-15
Sequence 15, Application US/09293238B
Patent No. 6455042
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
TITLE OF INVENTION: of Crohn's Disease by Administering an Antibody to Alpha R
TITLE OF INVENTION: Beta 7 Integrin
FILE REFERENCE: L0560/7005/ERP
CURRENT APPLICATION NUMBER: US/09/293,238B
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: US 08/879,338
PRIOR FILING DATE: 1997-06-20
PRIOR APPLICATION NUMBER: US 08/863,731
PRIOR FILING DATE: 1996-06-14
PRIOR APPLICATION NUMBER: US 08/199,776
PRIOR FILING DATE: 1994-02-18
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapien
US-09-293-238B 15

Query Match 2.1% Score 6; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 218 EDEEEE 223
IIIII
DB 5 EDEEEE 10

RESULT 13

US-09-293-238B-16
Sequence 16, Application US/09293238B
Patent No. 6455042
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
TITLE OF INVENTION: of Crohn's Disease by Administering an Antibody to Alpha R
TITLE OF INVENTION: Beta 7 Integrin
FILE REFERENCE: L0560/7005/ERP
CURRENT APPLICATION NUMBER: US/09/293,238B
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: US 08/879,338
PRIOR FILING DATE: 1997-06-20
PRIOR APPLICATION NUMBER: US 08/863,731
PRIOR FILING DATE: 1996-06-14
PRIOR APPLICATION NUMBER: US 08/199,776

PRIOR FILING DATE: 1994-02-18
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 16
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapien
US-09-293-238B-16

Query Match 2.1% Score 6; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 218 EDEEEE 223
IIIII
DB 4 EDEEEE 9

RESULT 14

PCT-US95-02044-15
Sequence 15, Application PC/TUS9502044
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Novel integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02044
FILING DATE: herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE: 18 February 1994
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
PCT-US95-02044-15

Query Match 2.1% Score 6; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 218 EDEEEE 223
IIIII
DB 5 EDEEEE 10


```

RESULT 15
PC1-US95-02044-16
: Sequence 16, Application PC/TUS9502044
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Novel integrin alpha subunit
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PC1-US95-02044
: FILING DATE: herewith
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/199,776
: FILING DATE: 18 February 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Plumer, Elizabeth R.
: REGISTRATION NUMBER: 36,637
: REFERENCE/DOCKET NUMBER: B0801/7020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 10 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: YES
: ANTI-SENSE: NO
: FRAGMENT TYPE: internal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: PC1-US95-02044-16

```

```

Query Match 2.1%; Score 6; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 218 EDEEE 223

Db 4 EDEEE 9

```

RESULT 16
US-08-199-776-22
: Sequence 22, Application US/08199776
: Patent No. 5594120
: GENERAL INFORMATION:
: APPLICANT: Brenner, Michael B.
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: No. 5594120e1 integrin alpha subunit
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/199,776
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Plumer, Elizabeth R.
: REGISTRATION NUMBER: 36,637
: REFERENCE/DOCKET NUMBER: B0801/7020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: YES
: ANTI-SENSE: YES
: ORIGINAL SOURCE:
: ORGANISM: synthetic peptide
: US-08-199-776-22
: Query Match 2.1%; Score 6; DB 1; Length 12;
: Best Local Similarity 100.0%; Pred. No. 30;
: Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
: Qy 218 EDEEE 223
: Db 6 EDEEE 11
: RESULT 17
: US-08-199-776-24
: Sequence 24, Application US/08199776
: Patent No. 5594120
: GENERAL INFORMATION:
: APPLICANT: Brenner, Michael B.
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: No. 5594120e1 integrin alpha subunit
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/199,776
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Plumer, Elizabeth R.
: REGISTRATION NUMBER: 36,637
: REFERENCE/DOCKET NUMBER: B0801/7020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12 amino acids

```

```
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? HYPOTHETICAL: YES
? ANTI-SENSE: NO
? ORIGINAL SOURCE: synthetic peptide
? ORGANISM: synthetic peptide
US-08-199-776-24

Query Match 2.1%; Score 6; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
Db 1 EDEEE 6

RESULT 18
US-08-663-731-22
: Sequence 22, Application US/08663731
: Patent No. 6057423
: GENERAL INFORMATION:
: APPLICANT: Brenner, Michael B.
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: No. 6057423el integrin alpha subunit
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/663,731
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/199,776
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Plummer, Elizabeth R.
: REGISTRATION NUMBER: 36,637
: REFERENCE/DOCKET NUMBER: B0801/7020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: YES
: ANTI-SENSE: NO
: ORIGINAL SOURCE: synthetic peptide
: ORGANISM: synthetic peptide
US-08-663-731-22

Query Match 2.1%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
Db 1 EDEEE 6

RESULT 20
US-08-879-338-22
: Sequence 22, Application US/08879338A
: Patent No. 6063906
: GENERAL INFORMATION:
: APPLICANT: Brenner, Michael B.
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/663,731
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/199,776
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Plummer, Elizabeth R.
: REGISTRATION NUMBER: 36,637
: REFERENCE/DOCKET NUMBER: B0801/7020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: YES
: ANTI-SENSE: NO
: ORIGINAL SOURCE: synthetic peptide
: ORGANISM: synthetic peptide
US-08-663-731-22

Query Match 2.1%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
Db 1 EDEEE 6
```

```
Db 6 EDEEE 11

RESULT 19
US-08-663-731-24
: Sequence 24, Application US/08663731
: Patent No. 6057423
: GENERAL INFORMATION:
: APPLICANT: Brenner, Michael B.
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: No. 6057423el integrin alpha subunit
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/663,731
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/199,776
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Plummer, Elizabeth R.
: REGISTRATION NUMBER: 36,637
: REFERENCE/DOCKET NUMBER: B0801/7020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: YES
: ANTI-SENSE: NO
: ORIGINAL SOURCE: synthetic peptide
: ORGANISM: synthetic peptide
US-08-663-731-24

Query Match 2.1%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
Db 1 EDEEE 6

RESULT 20
US-08-879-338-22
: Sequence 22, Application US/08879338A
: Patent No. 6063906
: GENERAL INFORMATION:
: APPLICANT: Brenner, Michael B.
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/663,731
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/199,776
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Plummer, Elizabeth R.
: REGISTRATION NUMBER: 36,637
: REFERENCE/DOCKET NUMBER: B0801/7020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: YES
: ANTI-SENSE: NO
: ORIGINAL SOURCE: synthetic peptide
: ORGANISM: synthetic peptide
US-08-663-731-24

Query Match 2.1%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
Db 1 EDEEE 6
```

; EARLIER FILING DATE: 1996-06-14
 ; PRIOR APPLICATION NUMBER: US 08/199,776
 ; PRIOR FILING DATE: 1994-02-18
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 12
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Peptide
 US-08-879-338-22

Query Match 2.1%; Score 6; DB 3; Length 12;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEF 223
 Db 6 EDEEEF 11

RESULT 22
 US-08-879-338-24
 ; Sequence 22, Application US/08879338A
 ; Patent No. 6063906
 ; GENERAL INFORMATION:
 ; APPLICANT: Brenner, Michael B.
 ; TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
 ; FILE REFERENCE: B0801/7080/ERP
 ; CURRENT APPLICATION NUMBER: US/08/879,338A
 ; EARLIER FILING DATE: 1997-06-20
 ; EARLIER APPLICATION NUMBER: US 08/663,731
 ; EARLIER FILING DATE: 1996-06-14
 ; EARLIER APPLICATION NUMBER: US 08/199,776
 ; PRIOR FILING DATE: 1994-02-18
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 24
 ; LENGTH: 12
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Peptide
 US-08-879-338-24

Query Match 2.1%; Score 6; DB 3; Length 12;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEF 223
 Db 1 EDEEEF 6

RESULT 22
 US-09-293-238B-22
 ; Sequence 22, Application US/09293238B
 ; Patent No. 6455042
 ; GENERAL INFORMATION:
 ; APPLICANT: Brenner, Michael B.
 ; TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
 ; FILE REFERENCE: L0560/7005/ERP
 ; CURRENT APPLICATION NUMBER: US/09/293,238B
 ; CURRENT FILING DATE: 1999-04-16
 ; PRIOR APPLICATION NUMBER: US 08/879,338
 ; PRIOR FILING DATE: 1997-06-20
 ; PRIOR APPLICATION NUMBER: US 08/663,731

; PRIOR FILING DATE: 1996-06-14
 ; PRIOR APPLICATION NUMBER: US 08/199,776
 ; PRIOR FILING DATE: 1994-02-18
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 12
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Peptide
 US-09-293-238B-22

Query Match 2.1%; Score 6; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEF 223
 Db 6 EDEEEF 11

RESULT 23
 US-09-293-238B-24
 ; Sequence 24, Application US/09293238B
 ; Patent No. 6455042
 ; GENERAL INFORMATION:
 ; APPLICANT: Brenner, Michael B.
 ; TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
 ; FILE REFERENCE: L0560/7005/ERP
 ; CURRENT APPLICATION NUMBER: US/09/293,238B
 ; CURRENT FILING DATE: 1999-04-16
 ; PRIOR APPLICATION NUMBER: US 08/879,338
 ; PRIOR FILING DATE: 1997-06-20
 ; PRIOR APPLICATION NUMBER: US 08/663,731
 ; PRIOR FILING DATE: 1996-06-14
 ; PRIOR APPLICATION NUMBER: US 08/199,776
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 24
 ; LENGTH: 12
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Peptide
 US-09-293-238B-24

Query Match 2.1%; Score 6; DB 4; Length 12;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEF 223
 Db 1 EDEEEF 6

RESULT 24
 PCT-US95-02044-22
 ; Sequence 22, Application PC/TUS9502044
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: Novel integrin alpha subunit
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
 ; STREET: 600 Atlantic Avenue
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA

ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02044
FILING DATE: herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE: 18 February 1994
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: YES
ORIGINAL SOURCE:
ORGANISM: synthetic peptide
PCT-US95-02044-22

Query Match 2.1% Score 6; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
Db 6 EDEEE 11

RESULT 25
PCT-US95-02044-24
Sequence 24, Application PC/US9502-44
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Novel integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02044
FILING DATE: herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE: 18 February 1994
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: synthetic peptide
PCT-US95-02044-24

Query Match 2.1% Score 6; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
Db 1 EDEEE 6

RESULT 26
US-08-199-776-18
Sequence 18, Application US/08199776
Patent No. 5594120
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 5594120el integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/199,776
FILING DATE:

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-199-776-18

Query Match 2.1% Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 218 EDEEEE 223
      |||||
Db 8 EDEEEE 13

RESULT 27
US-08-663-731-18
; Sequence 18, Application US/08663731
; Patent No. 6057423
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael B.
; APPLICANT: Parker, Christina M.
; TITLE OF INVENTION: No. 6057423el Integrin alpha subunit
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/663,731
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/199,776
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth K.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: 90801/7020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-663-731-18

Query Match 2.1% Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
      |||||
Db 8 EDEEEE 13

RESULT 28
US-08-879-338-18
; Sequence 18, Application US/08879338A
; Patent No. 6063906
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael B.
; APPLICANT: Parker, Christina M.
; TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
; SUBUNIT
```

```
; FILE REFERENCE: B0801/7080/ERP
; CURRENT APPLICATION NUMBER: US/08/879,338A
; CURRENT FILING DATE: 1997-06-20
; EARLIER APPLICATION NUMBER: US 08/663,731
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: US 08/199,776
; EARLIER FILING DATE: 1994-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-879-338-18

Query Match 2.1% Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
      |||||
Db 8 EDEEEE 13

RESULT 29
US-09-293-238B-18
; Sequence 18, Application US/09293238B
; Patent No. 6455042
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael B.
; APPLICANT: Parker, Christina M.
; TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
; TITLE OF INVENTION: or Crohn's Disease by Administering an Antibody to Alpha E
; FILE REFERENCE: L0560/7005/HRP
; CURRENT APPLICATION NUMBER: US/09/293,238B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: US 08/879,338
; PRIOR FILING DATE: 1997-06-20
; PRIOR APPLICATION NUMBER: US 08/663,731
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: US 08/199,776
; PRIOR FILING DATE: 1994-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-293-238B-18

Query Match 2.1% Score 6; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
      |||||
Db 8 EDEEEE 13

RESULT 30
PCT-US95-02044-18
; Sequence 18, Application PC/TUS9502044
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel integrin alpha subunit
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
```

```

? ZIP: 02210
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/02044
? FILING DATE: Herewith
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/199,776
? FILING DATE: 18 February 1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Plummer, Elizabeth R.
? REGISTRATION NUMBER: 36,637
? REFERENCE/DOCKET NUMBER: B0801/7020
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-720-3500
? TELEFAX: 617-720-2441
? INFORMATION FOR SEQ ID NO: 18:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 14 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? HYPOTHEICAL: YES
? ANTI-SENSE: NO
? FRAGMENT TYPE: internal
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? PC-US95-02044-18

Query Match      2.18; Score 6; DB 5; length 14;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  218 EDEEE 223
DB      8 EDEEE 13

RESULT 31
US-08-199-776 21
Sequence 21, Application US/08199776
Patent No. 5594120
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 5594120e: Integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Plummer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TELECOMMUNICATION INFORMATION:

```

```

? ZIP: 02210
? TELEPHONE: 617-720-3500
? TELEFAX: 617-720-2441
? INFORMATION FOR SEQ ID NO: 21:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 20 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? HYPOTHEICAL: YES
? ANTI-SENSE: NO
? FRAGMENT TYPE: internal
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? US-08-199-776-21

Query Match      2.18; Score 6; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  218 EDEEE 223
DB      14 EDEEE 19

RESULT 32
US-08-663-731-21
Sequence 21, Application US/08663731
Patent No. 6057423
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 6057423e: Integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663.731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Plummer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-663-731-21

```

```

Query Match          2.1%: Score 6; DB 3; Length 20;
Best Local Similarity 100.0%: Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
DB 14 EDEEE 19

RESULT 33
US-09-293 238B-21
: Sequence 21, Application: US/05293238B
: Patent No. 6455042
: GENERAL INFORMATION:
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
: FILE REFERENCE: 10560/7005/ERP
: CURRENT FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: US 08/879,338
: PRIOR FILING DATE: 1997-06-20
: PRIOR APPLICATION NUMBER: US 08/663,731
: PRIOR FILING DATE: 1996-06-14
: PRIOR APPLICATION NUMBER: US 08/199,776
: PRIOR FILING DATE: 1994-02-18
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 21
: LENGTH: 20
: TYPE: PRT
: ORGANISM: Homo sapiens
US 08-879-338-21

Query Match          2.1%: Score 6; DB 3; Length 20;
Best Local Similarity 100.0%: Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
DB 14 EDEEE 19

RESULT 34
US-09-293 238B-21
: Sequence 21, Application: US/05293238B
: Patent No. 6455042
: GENERAL INFORMATION:
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
: FILE REFERENCE: 10560/7005/ERP
: CURRENT FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: US 08/879,338
: PRIOR FILING DATE: 1997-06-20
: PRIOR APPLICATION NUMBER: US 08/663,731
: PRIOR FILING DATE: 1996-06-14
: PRIOR APPLICATION NUMBER: US 08/199,776
: PRIOR FILING DATE: 1994-02-18
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 21
: LENGTH: 20
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-293-238B-21

Query Match          2.1%: Score 6; DB 4; Length 20;
Best Local Similarity 100.0%: Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 218 EDEEE 223
DB 14 EDEEE 19

RESULT 35
PCT-US95-02044-21
: Sequence 21, Application PCT/TUS9502044
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Novel integrin alpha subunit
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/02044
: FILING DATE: herewith
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/199,776
: FILING DATE: 18 February 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Plumer, Elizabeth R.
: REGISTRATION NUMBER: 36,637
: REFERENCE/DOCKET NUMBER: B0801/7020
: IFTECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 21:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 20 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHETICAL: YES
: ANTI-SENSE: NO
: FRAGMENT TYPE: internal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
PCT-US95-02044-21

Query Match          2.1%: Score 6; DB 5; Length 20;
Best Local Similarity 100.0%: Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
DB 14 EDEEE 19

RESULT 36
US-08-199-776-5
: Sequence 5, Application US/08199776
: Patent No. 5594120
: GENERAL INFORMATION:
: APPLICANT: Brenner, Michael B.
: APPLICANT: Parker, Christina M.
: TITLE OF INVENTION: NO. 5594120e1 integrin alpha subunit
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield and Sacks, P.C.

```

STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/199,776
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: H0801/7020
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-199-776-5

Query Match 2.1%; Score 6; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
DB 14 EDEEEE 19
|||||

RESULT 37
US-08-663-731-5
Sequence 5, Application US/08663731
Patent No. 6057423
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 6057423el Integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESS: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: H0801/7020
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-663-731-5

Query Match 2.1%; Score 6; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
DB 14 EDEEEE 19
|||||

RESULT 38
US-08-879-338-5
Sequence 5, Application US/08879338A
Patent No. 6063906
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
TITLE OF INVENTION: Subunit
FILE REFERENCE: H0801/7080/ERP
CURRENT APPLICATION NUMBER: US/08/879,338A
CURRENT FILING DATE: 1997-06-20
EARLIER APPLICATION NUMBER: US 08/663,731
EARLIER FILING DATE: 1996-06-14
EARLIER APPLICATION NUMBER: US 08/199,776
EARLIER FILING DATE: 1994-02-18
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 21
TYPE: PRT
ORGANISM: Homo Sapiens
US-08-879-338-5

Query Match 2.1%; Score 6; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
DB 14 EDEEEE 19
|||||

RESULT 39
US-08-879-338-26
Sequence 26, Application US/08879338A
Patent No. 6063906
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: Antibodies to No. 6063906el Integrin Alpha
TITLE OF INVENTION: Subunit
FILE REFERENCE: H0801/7080/ERP
CURRENT APPLICATION NUMBER: US/08/879,338A


```

: CURRENT FILING DATE: 1997-06-28
: EARLIER APPLICATION NUMBER: US 08/663,731
: EARLIER FILING DATE: 1996-06-14
: EARLIER APPLICATION NUMBER: US 08/199,776
: EARLIER FILING DATE: 1994-02-18
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO: 26
: LENGTH: 21
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthetic peptide
US 08-879-338-26

```

```

Query Match      2.1%  Score 6;  DB 3;  Length 21;
Best Local Similarity 100.0%;  Pred. No. 50;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY  21B EDHEHE 223
    |||||
Db   14 EDEEEE 19

```

```

RESULT 4:
US-09-293-238R-5
: Sequence 5, Application US/09293238B
: Patent No. 6455042
: GENERAL INFORMATION:
: APPLICANT: Brenner, Michael R.
: TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
: TITLE OF INVENTION: or Crohn's Disease by Administering an Antibody to Alpha E
: FILE REFERENCE: Beta 7 Integritin
: FILE REFERENCE: L0560/7305/ERP
: CURRENT APPLICATION NUMBER: US/09/293,238B
: CURRENT FILING DATE: 1999-04-16
: PRIOR APPLICATION NUMBER: US 08/879,339
: PRIOR FILING DATE: 1997-06-20
: PRIOR APPLICATION NUMBER: US 08/663,731
: PRIOR FILING DATE: 1996-06-14
: PRIOR APPLICATION NUMBER: US 08/199,776
: PRIOR FILING DATE: 1994-02-18
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO: 5
: LENGTH: 21
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-293-238R-5

```

```

Query Match      2.1%  Score 6;  DB 4;  Length 21;
Best Local Similarity 100.0%;  Pred. No. 50;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

QY  21B EDHEHE 223
    |||||
Db   14 EDEEEE 19

```

```

Search completed: April 10, 2003, 10:42:57
Job time : 13.618 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:36:25 ; Search time 24.2188 Seconds
(without alignments)
2395.279 Million cell updates/sec

Title: US-09-930-026-2

Perfect score: 282

Sequence: 1 MAQQQMTSSQKALMLEKSL.....AMAPQPKPIHSGWGNTHSSC 282

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 671380 seqs, 20604715 residues

Word size : 0

Total number of hits satisfying chosen parameters: 60947

Minimum DB seq length: 0

Maximum DB seq length: 70

Post-processing: Listing first 135 summaries

Database :

SPTREMBL_21.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.protozoa.*
12: sp.virus.*
13: sp.vertibrate.*
14: sp.unclassified.*
15: sp.rv.*
16: sp.bacteriap.*
17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	2.8	59	10	Q38702
2	7	2.5	34	13	Q9PRE7
3	7	2.5	59	5	Q9UAT0
4	6	2.1	28	10	Q65311
5	6	2.1	40	8	Q312V4
6	6	2.1	41	10	Q96414
7	6	2.1	44	12	P87665
8	6	2.1	45	9	Q8W768
9	6	2.1	50	16	Q97SP9
10	6	2.1	51	11	Q62619
11	6	2.1	56	11	Q62618
12	6	2.1	57	3	P87196
13	6	2.1	57	16	Q98EA6
14	6	2.1	59	3	Q8TGW4
15	6	2.1	59	4	Q36Q74
16	6	2.1	62	7	Q9BCW0

Q9BCV9 saquinus oe
Q42130 arabidopsis
Q23725 chironomus
Q9A0G0 canis famil
Q54777 rattus norv
Q44850 borrelia bu
Q9R4X3 bacillus ce
Q9LWK7 trypanosoma
Q04817 sporobolus
Q9R4V8 streptococ
Q82084 human rhino
O11812 human immu
O11803 human immu
O11804 human immu
O11806 human immu
O11807 human immu
O11791 human immu
O11792 human immu
O11821 human immu
O11822 human immu
O11823 human immu
O11825 human immu
O11826 human immu
O11827 human immu
O11838 human immu
O11828 human immu
O11829 human immu
O11831 human immu
O11832 human immu
Q9W7Z6 human immu
Q9KAT7 bacillus ha
Q9S580 pseudomonas
Q9UC38 homo sapien
Q9XJ77 aureocumbra
Q9W800 human immu
Q9U6U9 plasmodium
Q9NFH6 plasmodium
Q9UEF5 homo sapien
Q9S940 homo sapien
Q9UMK5 homo sapien
Q9UM73 drosophila
Q9BM73 drosophila
O19501 homo sapien
Q9GHX3 hepatitis c
Q9IK15 hepatitis c
Q9HBQ2 homo sapien
Q9Y453 homo sapien
Q34897 lasiorhinus
Q34903 lasiorhinus
Q9UMQ3 homo sapien
Q9S658 homo sapien
Q9UD38 homo sapien
Q9UCF9 homo sapien
Q9VG96 rattus sp.
Q9V778 arabidopsis
Q9ZQ38 chlamydia t
Q9F6K1 nostoc sp.
Q9HBG1 homo sapien
Q9KXK7 mus musculus
Q50877 borrelia bu
Q11587 human immu
Q76310 human immu
Q13566 homo sapien
Q9M4D0 hordeum vul
Q8QZ2 simian viru
Q88774 rattus norv
Q9W806 human immu
Q05422 mycobacteri
Q26343 trypanosoma
Q8C3J8 drosophila
Q9Gau0 rana sylvat
Q96209 bacterioph
Q8I405 hepatitis c

90 5 1.8 38 15 Q9KMR9
 91 39 4 Q8WHR8
 92 5 2.8 40 8 Q34216
 93 5 1.8 40 5 Q9W729
 94 5 1.8 41 5 Q9W726
 95 5 1.8 41 9 Q9T156
 96 5 2.8 42 11 Q9KQF6
 97 5 1.8 42 2 P74973
 98 5 1.8 42 2 Q49078
 99 5 1.8 42 5 Q9N210
 100 5 1.8 42 8 Q34516
 101 5 1.8 42 10 Q9X1T4
 102 5 1.8 42 10 Q9C7M9
 103 5 1.8 43 4 Q9NFX5
 104 5 1.8 43 5 Q9XSK9
 105 5 1.8 43 10 Q9SQF2
 106 5 1.8 43 10 Q9SQD2
 107 5 1.8 43 10 Q9SQ01
 108 5 1.8 43 10 Q9SQD0
 109 5 1.8 43 10 Q9SQH4
 110 5 1.8 43 10 Q9SQA4
 111 5 1.8 43 10 Q9SCA0
 112 5 1.8 43 10 Q9SC88
 113 5 1.8 43 12 P87878
 114 5 1.8 43 16 Q8U5D4
 115 5 1.8 44 8 Q9GA18
 116 5 1.8 44 10 Q03996
 117 5 1.8 44 12 Q88417
 118 5 1.8 44 25 Q89363
 119 5 1.8 44 17 Q9RSG9
 120 5 1.8 45 8 Q12929
 121 5 1.8 45 10 Q3SA16
 122 5 1.8 45 12 Q91Q35
 123 5 1.8 45 12 Q91Q03
 124 5 1.8 45 27 Q82VW4
 125 5 1.8 47 4 Q81CF7
 126 5 1.8 47 8 Q9GA03
 127 5 1.8 47 15 Q9W7Y8
 128 5 1.8 47 15 Q8Q399
 129 5 1.8 48 2 Q9EYX3
 130 5 1.8 48 4 Q9P151
 131 5 1.8 48 5 Q9VR85
 132 5 1.8 48 5 Q9TXD4
 133 5 1.8 48 13 Q9XH50
 134 5 1.8 48 13 Q9XH49
 135 5 1.8 48 13 Q9YH48

ALIGNMENTS

RESULT 1
 Q38702 PRELIMINARY: PRT: 59 AA.
 AC Q38702:
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-NOV-1996 (TEMBLrel. 08, Last annotation update)
 DE Nondormancy-associated clone AFN2 putative ORF1 (Fragment).
 OS Avena fatua.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 CC Avenae; Avena.
 OX NCBI_TaxID=4499;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AN265; TISSUE=EMBRIO;
 RX MEDLINE=95306780; PubMed=7781716;
 RA Johnson R.R., Cranston H.J., Chaverra M.E., Dyer W.F.;
 RI *Characterization of cDNA clones for differentially expressed genes in
 RI embryos of dormant and nondormant Avena fatua L. caryopses.;
 RI Plant Mol. Biol. 28:113-122(1995).
 DR FMBL; J19999; AAA76741.1; -.

FT NON_TER 1
 SQ SEQUENCE 59 AA; 6549 MW; 4244310FH427FEA6 CHC64;
 Query Match 2.8%; Score 8; DB 10; Length 59;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 216 DDEDEEE 223
 Db 43 DDEDEEE 50
 RESULT 2
 Q9PRE7 PRELIMINARY: PRT: 34 AA.
 AC Q9PRE7:
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE PAX-3 (Fragment).
 GN PAX-3.
 OS Oryzias latipes (Medaka fish).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neotocostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 CC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DA, AND HNT;
 RX MEDLINE=20082974; PubMed=10613850;
 RA Ohtsuka M., Makino S., Yoda K., Wada H., Naruse K., Mitani H.,
 RA Shima A., Ozato K., Kimura M., Inoko H.;
 RI *Construction of a linkage map of the Medaka (Oryzias latipes) and
 RI mapping of the Da mutant locus defective in dorsoventral patterning.;
 RI Genome Res. 9:1277-1287(1999).
 DR FMBL; AB030470; HAA85137.1; -.
 DR FMBL; AB030469; HAA85136.1; -.
 DR InterPro; IPR001523; Paired_box.
 DR Pfam; PF00292; PAX; 1.
 FT NON_TER 1
 SQ SEQUENCE 34 AA; 3737 MW; E42A3AC06AEFA48 CHC64;
 Query Match 2.5%; Score 7; DB 13; Length 34;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 219 EDEDEEE 224
 Db 22 EDEDEEE 28
 RESULT 3
 Q9UAT0 PRELIMINARY: PRT: 59 AA.
 AC Q9UAT0:
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE Glucose-6-phosphate 1-dehydrogenase (Fragment).
 GN ZW.
 OS Ceratitis capitata (Mediterranean fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Tephritidae; Tephritidae; Ceratitis.
 OX NCBI_TaxID=7213;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HLAB;
 RX MEDLINE=99376744; PubMed=1047886;
 RA He M., Haymer D.S.;
 RI *Genetic relationships of populations and the origins of new

RT infestations of the Mediterranean fruit fly.*;
 RI MOI: 8:1247-1257(1999).
 DR EMBL: AF125159; AAD32569.1; -;
 LR InterPro: IPR001282; G6PD.
 DR Pfam: PF00479; G6PD; 1.
 DR ProDom: PD001129; G6PD; 1.
 FT NON_TER 59 1
 FT NON_TER 59 1
 SQ SEQUENCE 59 AA; 6966 MW; 9204GCD1D2AF6706 CRC64;

Query Match 2.5%; Score 7; DB 5; Length 59;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 GASGDLA 263

Db 7 GASGDLA 13

RESULT 4

ID 065311 PRELIMINARY; PRT; 28 AA.

AC 065311;

DT 01-AGS-1998 (TRENBLrel. 07, Created)

DI 01-AGS-1998 (TRENBLrel. 07, Last sequence update)

DR 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE Beta-9 tubulin (Fragment).

GN TUB9.

OS Arabidopsis thaliana (Mouse ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID:3702;

RN 1;

RN 1;

RP SEQUENCE FROM N.A.

RC STRAIN-CV. WS;

RA Sanders P.M., Rai A.O., Weterings K., McIntire K.N., Hsu Y.C.,

Lee P.Y., Irong M.T., Heals T.H., Goldberg R.A.;

RT "Archer Development Defects in Arabidopsis thaliana Male-Sterile

Mutants.";

KL Sex. Plant Reprod. 11:297-322(1999).

DR EMBL: AF060248; AAC97107.1; -;

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 28 AA; 3388 MW; 8060E4E537BE670E CRC64;

Query Match 2.1%; Score 6; DB 10; Length 28;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223

Db 21 EDEEE 26

RESULT 5

Q912V4

ID 0912V4 PRELIMINARY; PRT; 40 AA.

AC 0912V4;

DT 01-MAY-2000 (TRENBLrel. 13, Created)

DI 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DR 01-JUN-2000 (TRENBLrel. 14, Last annotation update)

DE Cytochrome c oxidase subunit 7 (Fragment).

OS Crithidia fasciculata.

OC Mitochondrion.

OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.

OX NCBI_TaxID:5656;

RN 1;

RN 1;

RP SEQUENCE

EX MEDLINE-97001679; PubMed-8844672;

RA Speijer D., Mulijers A.O., Dekker E., de Haan A., Breek C.K.,

Albracht S.P., Benne R.;

RT "Purification and characterization of cytochrome c oxidase from the

insect trypanosomatid Crithidia fasciculata.";

RL Mol. Biochem. Parasitol. 79:47-59(1996).
 SQ SEQUENCE 40 AA; 4801 MW; 265C172724820F84 CRC64;

Query Match 2.1%; Score 6; DB 8; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 TTLAEY 190

Db 11 TTLAEY 16

RESULT 6

Q96414

ID 096414 PRELIMINARY; PRT; 41 AA.

AC 096414;

DT 01-FEB-1997 (TRENBLrel. 02, Created)

DI 01-FEB-1997 (TRENBLrel. 02, Last sequence update)

DR 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DR Beta-tubulin 3 (Fragment).

OS Daucus carota (Carrot).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.

OX NCBI_TaxID:4039;

RN 1;

RN 1;

RP SEQUENCE FROM N.A.

RC STRAIN-CV. KINTOKI;

RA Okamura S., Sonehara K., Naito K., Ohkawa H., Kuramori S., Talsuta M.,

Minamizono M.;

RT "Characterization of beta-tubulin genes of carrot.";

RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U64430; AAB47936.1; -;

DR InterPro: IPR003008; Tubulin_FtsZ.

DR Pfam: PF00691; tubulin; 1.

KW GTP-binding.

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 41 AA; 4945 MW; A5236AAE45F10274 CRC64;

Query Match 2.1%; Score 6; DB 10; Length 41;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223

Db 32 EDEEE 37

RESULT 7

P87665

ID P87665 PRELIMINARY; PRT; 44 AA.

AC P87665;

DT 01-MAY-1997 (TRENBLrel. 03, Created)

DI 01-DEC-2001 (TRENBLrel. 19, Last sequence update)

DR 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE DNA-binding protein (Fragment).

GN E2A DBP.

OS duck adenovirus 1.

OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Adenovirus.

OX NCBI_TaxID:130329;

RN 1;

RN 1;

RP SEQUENCE FROM N.A.

RC STRAIN-NORTHERN IRELAND ISOLATE VF76-127;

RA Harrach B., Meehan B.M., Benko M., Adair B.M., Todd D.;

RT "Close phylogenetic relationship between egg drop syndrome virus,

RT bovin adenovirus serotype 7, and ovine adenovirus strain 287.";

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U63515; AAB39492.1; -;

KW DNA-binding.

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 44 AA; 5065 MW; 940EB759FA638612 CRC64;

Query Match 2.1%; Score 6; DB 12; Length 44;

Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 DEDEED 222
11, 111
DB 36 DEDEEE 41

RESULT 8

Q8W768 ID Q8W768 PRELIMINARY; PRT; 45 AA.
AC Q8W768;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Bacteriophage P22 aa - cal genes.
OS Bacteriophage P22.
GC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC P22-like viruses;
OX NCBI_TaxID=10754;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81080981;
RA Youderian P., Susskind M.M.;
KT "Bacteriophage P22 proteins specified by the region between genes 9 and 11";
RI Virology 107:270-282(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82247245;
RA Storme G.D., Schneider I.D., Gold L., Ehrenfeucht A.;
RI "Use of the 'perception' algorithm to distinguish translational initiation sites in E. coli.";
RL Nucleic Acids Res. 10:2997-3011(1982).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82247244;
RA Storme G.D., Schneider I.D., Gold L.M.;
KT "Characterization of translational initiation sites in E. coli.";
RI Nucleic Acids Res. 10:2971-2996(1982).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93223617;
RA Benedik M., Mascarenhas D., Campbell A.;
KT "The integrase promoter and 5' terminator in bacteriophages Lambda and 434.";
RI Virology 126:658-668(1983).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=87060955; PubMed=3491212;
RA Leong J.M., Nunes-baby S.E., Oser A.B., Lesser C.F., Youderian P., Susskind M.M., Landy A.;
KT "Structural and regulatory divergence among site-specific recombination genes of lambdaoid phage.";
RI J. Mol. Biol. 189:603-616(1986).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=88013195; PubMed=3660589;
RA Murphy K.C., Fenton A.C., Poteete A.R.;
KT "Sequence of the bacteriophage P22 anti-recBCD (abc) genes and properties of P22 abc region deletion mutants.";
RI Virology 160:456-464(1987).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=94018622; PubMed=8412679;
RA Walli D.L., Ho Y.S., Powers S., Rosenberg M.;
RI "The int genes of bacteriophages P22 and lambda are regulated by different mechanisms.";
RL Mol. Microbiol. 9:261-271(1993).
DR EMBL; L06296; AAC18886.1;
SO SEQUENCE 45 AA; 5082 MW; 8110C925E27F44DF CRC64;

Query Match 2.1%; Score 6; DB 9; Length 45;

Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GGYFKA 60
111111
DB 20 GGYFKA 25

RESULT 9

Q97SP9 ID Q97SP9 PRELIMINARY; PRT; 50 AA.
AC Q97SP9;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical protein SP0277.
CN SP0277.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D., Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J., Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D., Maynam L.A., White O., Salzberg S.L., Lewis M.R., Radune D., Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L., McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.G., Venter J.C., Dougherty H.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
KT "Complete genome sequence of a virulent isolate of Streptococcus pneumoniae";
RI Science 293:498-506(2001).
RL EMBL; AE007341; AAK74455.1;
DR TIGR; SP0277;
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 50 AA; 5524 MW; 9B1E5A4DBF3477A CRC64;

Query Match 2.1%; Score 6; DB 16; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 RTILIS 126
111111
DB 25 RTILIS 30

RESULT 10
Q62619 ID Q62619 PRELIMINARY; PRT; 51 AA.
AC Q62619;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Fetal troponin T 3 (fragment).
CN TN3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKLELETAL MUSCLE;
RX MEDLINE=93345743; PubMed=8344466;
RA Briggs M.M., Schachar F.;
KT "Origin of fetal troponin T: Developmentally regulated splicing of a new exon in the fast troponin T gene.";
RI Dev. Biol. 158:503-509(1993).
RN [2]
RP SEQUENCE FROM N.A.

RC TISSUE-SKELETAL MUSCLE;

RA Briggs M.M.;

RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL: J04960; AAL16033.1; -. POTENTIAL.

FT CHAIN 2 >53

FT NON_TER 51

SQ SEQUENCE 51 AA: 6113 MW: A4ACRA5B3PB6A25D CRC64;

Query Match 2.1%; Score 6; DB 11; Length 51;

Best Local Similarity 100.0%; Pred. No. 2.3e-02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 DEDEEE 223

Db 37 EDEEEE 42

RESULT 11

Q62618

ID Q62618 PRELIMINARY; PRT: 56 AA.

AC Q62618;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Actal troponin T 2 (Fragment).

GN TN2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

CX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-SKELETAL MUSCLE;

RX MEDLINE: 93345743; PubMed: 8344456;

RA Briggs M.M., Schachar F.;

RT "Origin of fetal troponin T: developmentally regulated splicing of a

RT new exon in the fast troponin T gene."

RL Dev. Biol. 138:503-508(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-SKELETAL MUSCLE;

RA Briggs M.M.;

RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL: U04979; AAL16032.1; -. POTENTIAL.

FT CHAIN 2 >56

FT NON_TER 56

SQ SEQUENCE 56 AA: 6728 MW: 99B1C3E69337E886 CRC64;

Query Match 2.1%; Score 6; DB 11; Length 56;

Best Local Similarity 100.0%; Pred. No. 2.5e-02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223

Db 42 EDEEEE 47

RESULT 12

P87196

ID P87196 PRELIMINARY; PRT: 57 AA.

AC P87196;

DT 01-JUN-1997 (TrEMBLrel. 04, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE NAP1 protein (Fragment).

GN NAP1 OR YKR048C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.

CX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RA Viissers S., Urrestarazu L.A., Jauniaux J.C.;

RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA MIPS;

RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z28272; CAA82124.2; -. POTENTIAL.

DR SCQ: S0001756; NAP1. 1

FT NON_TER 1

SQ SEQUENCE 57 AA: 6424 MW: F9DD2C1BCF0F5B33 CRC64;

Query Match 2.1%; Score 6; DB 3; Length 57;

Best Local Similarity 100.0%; Pred. No. 2.5e-02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 DEDEEE 222

Db 19 DEDEEE 24

RESULT 13

Q98EA6

ID Q98EA6 PRELIMINARY; PRT: 57 AA.

AC Q98EA6;

DT 01-OCT-2001 (TrEMBLrel. 18, Created)

DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)

DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)

DE Thioredoxin-related CycX 3' region protein, inner membrane

DE protein.

GN MSL4331.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Phyllobacteriaceae; Mesorhizobium.

CX NCBI_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAF303099;

RX MEDLINE=21082930; PubMed=11214968;

RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shampo S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti."

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB51014.1; -. POTENTIAL.

KW Complete proteome.

SQ SEQUENCE 57 AA: 6141 MW: 62141B377E6ACAC4 CRC64;

Query Match 2.1%; Score 6; DB 16; Length 57;

Best Local Similarity 100.0%; Pred. No. 2.5e-02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 KAAAEK 178

Db 51 KAAAEK 56

RESULT 14

Q8TGW4

ID Q8TGW4 PRELIMINARY; PRT: 59 AA.

AC Q8TGW4;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Aspartate aminotransferase (fragment).

DE Tuber borchii (White truffle).

OC Fukariyota; Fungi; Ascomycota; Pezizomycotina; Pezizomycetes;

OC Pezizales; Tuberales; Tuberaceae; Tuber.

CX NCBI_TaxID=42251;

RN [1]

RP SEQUENCE FROM N.A.

RA Montanini H., Ottonello S.;

```

RT *Characterization of an aspartate aminotransferase in the symbiotic
RI fungus Tuber borchii.
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF462031; AAL76244.1;
KW Transferase; Aminotransferase.
FT NON_TER 59 59
FT NON_TER 59 59
SQ SEQUENCE 59 AA: 6875 MW: 49607.077 DAE944E3 CRC64;

Query Match 2.1% Score 6; DB 3; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.6e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 258 ASGDLA 263
DB 32 ASGDLA 37

RESULT 15
Q96C74 ID Q96C74 PRELIMINARY: PRT: 59 AA.
AC Q96C74
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Mitochondrial ribosomal protein L7/L12 (Fragment).
GN MRPL7/L12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-21429115; PubMed-11543634;
RA Kennecott N., Suzuki T., Uechi T., Maqsoori M., Kuniba M., Higa S.,
RW Watanabe K., Tanaka T.;
RT The human mitochondrial ribosomal protein genes: Mapping of 54 genes
RL to the chromosomes and implications for human disorders.
RL Genomics 77:65-75(2001).
DR EMBL: AB051337; BAB54927.1;
DR InterPro: IPRC00206; Ribosomal_L12.
DR Pfam: PF00542; Ribosomal_L12; 1.
DR ProDom: PD001326; Ribosomal_L12; 1.
KW Ribosomal protein.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 59 AA: 6385 MW: 8C6933B8E88B4100 CRC64;

Query Match 2.1% Score 6; DB 4; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.6e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 KAEAEK 178
DB 39 KAEAEK 44

RESULT 16
Q9BCW0 ID Q9BCW0 PRELIMINARY: PRT: 62 AA.
AC Q9BCW0
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II antigen (Fragment).
OS Saguinus oedipus (Cotton-top tamarin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
CX NCBI_TaxID:9490;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-249SAGE-DOA-I;
RW Kriener K., Ohtsuga C., Klein J.;
RT *Independent origin of functional MHC class II genes in humans and new

```

```

RT *Independent origin of functional MHC class II genes in humans and new
RI world monkeys.
RL Hum. Immunol. 62:1-14(2001).
DR EMBL: AY013373; AAK00533.1;
DR InterPro: IPR003597; Ig-CL.
DR InterPro: IPR003036; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00407; IGcl; 1.
FT NON_TER 1 1
FT NON_TER 62 62
SQ SEQUENCE 62 AA: 6707 MW: 8EFCFDC244E185B1 CRC64;

Query Match 2.1% Score 6; DB 7; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.7e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GLPNTL 52
DB 4 GLPNTL 9

RESULT 17
Q9BCV9 ID Q9BCV9 PRELIMINARY: PRT: 52 AA.
AC Q9BCV9
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II antigen (Fragment).
OS Saguinus oedipus (Cotton-top tamarin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
CX NCBI_TaxID:9490;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-249SAGE-DOA-I;
RW Kriener K., Ohtsuga C., Klein J.;
RT *Independent origin of functional MHC class II genes in humans and new
RI world monkeys.
RL Hum. Immunol. 62:1-14(2001).
DR EMBL: AY013373; AAK00533.1;
DR InterPro: IPR003597; Ig-CL.
DR InterPro: IPR003036; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00407; IGcl; 1.
FT NON_TER 1 1
FT NON_TER 62 62
SQ SEQUENCE 62 AA: 6759 MW: 8EFCF8D564C4385B1 CRC64;

Query Match 2.1% Score 6; DB 7; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.7e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 GLPNTL 52
DB 4 GLPNTL 9

RESULT 18
Q42130 ID Q42130 PRELIMINARY: PRT: 64 AA.
AC Q42130
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Guanine nucleotide-binding protein (fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
CX NCBI_TaxID:3702;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN CV. COLUMBIA: TISSUE-SEEDLING;
RA Hoste H.;
RI Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
RR EMBL: Z26552; CAA8323.1;
DR InterPro: IPR001580; WD40.
DK Pfam: PF03430; WD40.1;
DR SMART: SM03320; WD40.1;
DR PROSITE: PS0082; WD_REPEATS.2; 1.
DR PROSITE: PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT NON_TER 1
SC SEQUENCE 64 AA: 6945 MW: 573564192287C993 CRC64;

Query Match 2.1%; Score 6; DB 10; Length 64;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 KAAEK 178
Db 14 KAAEK 19

RESULT 19
Q23725 PRELIMINARY: PRI: 69 AA.
AC Q23725;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Gamma protein constant region (Fragment).
GN HRI.
OS Chironomus pallidivittatus (Midge).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironomidae; Chironomus.
OX NCBI_TaxID=7151;
RN 1;
RP SEQUENCE FROM N.A.
RA Lendahl G., Salgia H., Hoozee C., Edstrom J.E., Wieslander L.;
RI "Rapid and Concerted Evolution of Repeat Units in a Balbiani Ring
RT Gene";
RL Genetics 117:43-49(1987).
DR EMBL: X06433; CAA2939.1;
FT NON_TER 69
FT NON_TER 69
SC SEQUENCE 69 AA: 7216 MW: 16FA25C6A5644B48 CRC64;

Query Match 2.1%; Score 6; DB 5; Length 69;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 KAAEK 178
Db 59 KAAEK 64

RESULT 20
Q23725 PRELIMINARY: PRI: 15 AA.
AC Q23725;
DT 01-OCT-2003 (TREMBlrel. 15, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Class 2 beta-tubulin (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN 1;
RP SEQUENCE FROM N.A.
RA Atal K.;
RI "Molecular cloning of isotype-specific regions of five classes of

```

```

RT canine beta-tubulin and their tissue distribution.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB02054; BAA36409.3;
KW GTP-binding.
FT NON_TER 1
FT NON_TER 15
SC SEQUENCE 15 AA: 1713 MW: 20511A5F599F5A1D CRC64;

Query Match 1.8%; Score 5; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 EEED 224
Db 2 EEED 6

RESULT 21
Q54777 PRELIMINARY: PRI: 17 AA.
AC Q54777;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE PSD-95 binding protein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE-98089008; PubMed-9428732;
RA Kawashima N., Takamiya K., Sun J., Kitabatake A., Schue K.;
RI "Differential expression of isoforms of PSD-95 binding protein
RT (KAP/SAPAP) during rat brain development.";
RL FEBS Lett. 418:301-304(1997).
DR EMBL: AH005146; BAA24285.1;
FT NON_TER 17
FT NON_TER 17
SC SEQUENCE 17 AA: 1934 MW: B74991395C4C0AE7 CRC64;

Query Match 1.8%; Score 5; DB 11; Length 17;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 VSATK 173
Db 13 VSATK 17

RESULT 22
Q44850 PRELIMINARY: PRI: 19 AA.
AC Q44850;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical 2.3 kDa protein (Fragment).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN-297;
RA Akins D.R., Popova T., Brusca J., Goldberg M.L., Li M., Baker S.C.;
RI "Use of PhA Gene Fusions and Anchored PCR to Identify and Express
RT Borrelia burgdorferi Candidate Outer Membrane Proteins.";
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: L31425; AAA64501.1;
KW Hypothetical protein.
FT NON_TER 19
FT NON_TER 19
SC SEQUENCE 19 AA: 2321 MW: C821BF312DB41512 CRC64;

```


Query Match 1.8% Score 5; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 SILHP 59
 DB 9 SILHP 13

RESULT 23
 Q9R4X3 PRELIMINARY; PRT; 19 AA.

AC Q9R4X3
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
 DE Amidase (Fragment).
 OS Bacillus cereus.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 CC Bacillaceae; Bacillus.
 OX NCBI_TaxID=1196;
 RN 1
 RP SEQUENCE
 RX MEDLINE=94362506; PubMed=8081503;
 RA Makino S., Ito N., Inoue T., Miyata S., Moriyama R.;
 RT "A spore lytic enzyme released from Bacillus cereus spores during
 germination."
 RL Microbiology 140:1403-1410(1994).
 SQ SEQUENCE 19 AA; 2103 MW; 279DAJ5BRC0C54D2 CRC64;

Query Match 1.8% Score 5; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 RGASG 260
 DB 8 RGASG 12

RESULT 24
 Q9TWK7 PRELIMINARY; PRT; 19 AA.

AC Q9TWK7
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
 DE Haptoglobin-related protein alpha subunit (fragment).
 OS Trypanosoma brucei.
 CC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN 1
 RP SEQUENCE
 RX MEDLINE=95232503; PubMed=7716520;
 RA Smith A.B., Esso J.D., Hajduk S.L.;
 RT "Killing of trypanosomes by the human haptoglobin-related protein."
 RL Science 268:284-286(1995).
 SQ SEQUENCE 19 AA; 2072 MW; 5B82440A3D8A8D6 CRC64;

Query Match 1.8% Score 5; DB 5; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 VISAL 131
 DB 5 VISAL 9

RESULT 25
 Q04817 PRELIMINARY; PRT; 19 AA.

AC Q04817
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Hypothetical 2.3 kDa protein (Fragment).
 OS Sporobolus stapfianus (Resurrection grass).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 CC Chloridoideae; Eragrostidae; Sporobolus.
 OX NCBI_TaxID=56623;
 RN 1
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Blomstedt C.K., Gianello R.D., Neale A.D., Hamill J.D., Gaff D.F.;
 RT "Isolation and characterization of cDNAs associated with the onset of
 desiccation tolerance in the resurrection grass Sporobolus
 stapfianus."
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y10781; CAA17531;
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 19 AA; 2270 MW; 4C644625813F244E CRC64;

Query Match 1.8% Score 5; DB 10; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 LYDDL 210
 DB 3 LYDDL 7

RESULT 26
 Q9R4V8 PRELIMINARY; PRT; 20 AA.

AC Q9R4V8
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
 DE Erythrogenic toxin type C (Fragment).
 OS Streptococcus pyogenes.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 CC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1314;
 RN 1
 RP SEQUENCE
 RX MEDLINE=95004281; PubMed=7920465;
 RA Ozegowski J.H., Wollweber L., Schmidt K.H., Vettermann S.,
 Reichardt W., Kohler W.;
 RT "Streptococcal erythrogenic toxin type C is not a phosphorylated
 protein. Description of two different purification procedures and
 investigation of its phosphorylation state."
 RL FEMS Immunol. Med. Microbiol. 9:65-76(1994).
 SQ SEQUENCE 20 AA; 2238 MW; A3E9A1AF496D4E26 CRC64;

Query Match 1.8% Score 5; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 SDLLY 207
 DB 10 SDLLY 14

RESULT 27
 Q82084 PRELIMINARY; PRT; 21 AA.

AC Q82084
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Type 1A (HRV-1A) genomic-linked viral protein (Vp) genomic region
 DE (Fragment).
 OS Human rhinovirus sp.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Rhinovirus.

```

OX NCBI_TaxID-169066;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE-96144022; PubMed-3512851;
RA Werner G., Rosenzweig B., Bauer E.A., Seifert J.-M., Werner F.-J.,
RA Besmer J.;
RT "Molecular cloning and sequence determination of genomic regions
RT encoding protease and genomic-linked protein of three
RT picornaviruses.";
RL J. Virol. 57:1084-1093(1986).
DR EMBL; M12166; AAA45760.1; -
FT NON_TER 1 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2324 MW; 5DCE71ED80B942 CRC64;

Query Match 1.8%; Score 5; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 KTKVP 197
DB 10 KTKVP 14
|||||

RESULT 28
O11812
ID O11812 PRELIMINARY; PRT; 21 AA.
AC O11812;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
GS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID-11676;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Coulter P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT "Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.";
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL; Y12118; CAA72818.1; -
FT NON_TER 1 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2412 MW; 054EA519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8
|||||

RESULT 29
O11803
ID O11803 PRELIMINARY; PRT; 21 AA.
AC O11803;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
GS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID-11676;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Coulter P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT "Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.";
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL; Y12118; CAA72818.1; -
FT NON_TER 1 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2412 MW; 054EA519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8
|||||

RESULT 29
O11803
ID O11803 PRELIMINARY; PRT; 21 AA.
AC O11803;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
GS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID-11676;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Coulter P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT "Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.";
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL; Y12118; CAA72818.1; -
FT NON_TER 1 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2412 MW; 054EA519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8
|||||

RESULT 30
O11804
ID O11804 PRELIMINARY; PRT; 21 AA.
AC O11804;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
GS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID-11676;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Coulter P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT "Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.";
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL; Y12118; CAA72818.1; -
FT NON_TER 1 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2412 MW; 054EA519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8
|||||

RESULT 31
O11806
ID O11806 PRELIMINARY; PRT; 21 AA.
AC O11806;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)

```

```

OX Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID-11676;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Coulter P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT "Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.";
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL; Y12118; CAA72818.1; -
FT NON_TER 1 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2412 MW; 0543F619H3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8
|||||

RESULT 30
O11804
ID O11804 PRELIMINARY; PRT; 21 AA.
AC O11804;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
GS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID-11676;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Coulter P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT "Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.";
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL; Y12118; CAA72818.1; -
FT NON_TER 1 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2412 MW; 055B4519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8
|||||

RESULT 31
O11806
ID O11806 PRELIMINARY; PRT; 21 AA.
AC O11806;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)

```

```

DI 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID:11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Goulder P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT *Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.*;
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL: Y12133; CAA72836.1;
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2426 MW; 0543E619B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8

RESULT 32
DI 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID:11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Goulder P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT *Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.*;
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL: Y12133; CAA72836.1;
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2426 MW; 0543E619B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8

```

```

RESULT 33
DI 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID:11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Goulder P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT *Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.*;
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL: Y12133; CAA72836.1;
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2412 MW; 055B4519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
DB 4 ELKSL 8

RESULT 34
DI 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID:11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-9126923;
RA Goulder P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT *Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.*;
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL: Y12133; CAA72836.1;
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2350 MW; 0543E619B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 16 ELKSL 20
 DB 4 ELKSL 8

RESULT 35

ID O11821 PRELIMINARY: PRT: 21 AA.
 AC O11821;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97272078; PubMed-9126923;
 RA Gould P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
 RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
 RA McMichael A.J.;
 RT *Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
 RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
 RT by epitope mutation.*;
 RL J. Exp. Med. 185:1423-1433(1997).
 DR EMBL: Y12147; CAA72846.1;
 DR InterPro: IPR000071; Retrovir_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 KW AIDS; Core protein; Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2394 MW; 5F5B4519AB673DB5 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
 DB 4 ELKSL 8

RESULT 36

ID O11822 PRELIMINARY: PRT: 21 AA.
 AC O11822;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97272078; PubMed-9126923;
 RA Gould P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
 RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
 RA McMichael A.J.;
 RT *Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
 RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
 RT by epitope mutation.*;
 RL J. Exp. Med. 185:1423-1433(1997).
 DR EMBL: Y12148; CAA72846.1;
 DR InterPro: IPR000071; Retrovir_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 KW AIDS; Core protein; Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2408 MW; 5F43E619AB673DB5 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
 DB 4 ELKSL 8

RESULT 37

ID O11823 PRELIMINARY: PRT: 21 AA.
 AC O11823;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97272078; PubMed-9126923;
 RA Gould P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
 RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
 RA McMichael A.J.;
 RT *Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
 RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
 RT by epitope mutation.*;
 RL J. Exp. Med. 185:1423-1433(1997).
 DR EMBL: Y12149; CAA72846.1;
 DR InterPro: IPR000071; Retrovir_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 KW AIDS; Core protein; Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2396 MW; 055B4519AB673DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
 DB 4 ELKSL 8

RESULT 38

ID O11825 PRELIMINARY: PRT: 21 AA.
 AC O11825;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97272078; PubMed-9126923;
 RA Gould P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
 RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
 RA McMichael A.J.;
 RT *Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
 RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
 RT by epitope mutation.*;
 RL J. Exp. Med. 185:1423-1433(1997).
 DR EMBL: Y12151; CAA72851.1;
 DR InterPro: IPR000071; Retrovir_p17.

```

DR Pfam: PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 21
SQ SEQUENCE 21 AA; 2412 MW; 055B4519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
   ||||
Db 4 ELKSL 8

RESULT 39
Q11826
ID Q11826 PRELIMINARY; PRT; 21 AA.
AC Q11826;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DI 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DI 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID:11676;
[]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed-3126923;
RA Gouldner P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT "Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte
R2 responses in two HLA-identical siblings with HLA-A*020: are influenced
R3 by epitope mutation.";
RL J. Exp. Med. 185:1423-433(1997).
DR EMBL: Y12153; CAA72852.1;
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 21
SQ SEQUENCE 21 AA; 2412 MW; 055B4519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
   ||||
Db 4 ELKSL 8

RESULT 40
Q11827
ID Q11827 PRELIMINARY; PRT; 21 AA.
AC Q11827;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DI 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DI 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
CX NCBI_TaxID:11676;
[]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE-97272078; PubMed 9126923;
RA Gouldner P.J.R., Sewell A.K., Lalloo D.G., Price D.A., Whelan J.A.,
RA Evans J., Taylor P., Luzzi G., Giangrande P.L., Phillips R.E.,
RA McMichael A.J.;
RT "Patterns of immunodominance in HIV-1-specific cytotoxic T lymphocyte

```

```

RT responses in two HLA-identical siblings with HLA-A*0201 are influenced
RT by epitope mutation.";
RL J. Exp. Med. 185:1423-1433(1997).
DR EMBL: Y12153; CAA72853.1;
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 21
SQ SEQUENCE 21 AA; 2412 MW; 055B4519B3C43DB6 CRC64;

Query Match 1.8%; Score 5; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELKSL 20
   ||||
Db 4 ELKSL 8

Search completed: April 10, 2003, 10:41:11
Job time : 29.2188 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:33:55 ; Search time 7.29882 Seconds
(without alignments)
1602.494 Million cell updates/sec

Title: US-09-930-026-2
Perfect score: 282
Sequence: 1 MAQQQTSSQKALMLEKSL.....ANAPQKPIHSGWNTSSC 282

Scoring table: OIIGO
Gapop 60.0 , Gapext 60.0

Searched: 112802 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6866

Minimum DB seq length: 0
Maximum DB seq length: 70

Post-processing: Listing first 135 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	2.1	21	1 F1H8_CEREL	P14468 cervus elap
2	6	2.1	29	1 HS98_NEUCR	P31540 neurospora
3	6	2.1	38	1 YF07_HAEIN	P44229 haemophilus
4	6	2.1	46	1 A0FA_SHEEP	O18851 ovis aries
5	6	2.1	51	1 YF93_HAEIN	P44264 haemophilus
6	6	2.1	65	1 PATP_PHYAM	P81418 phytolacca
7	6	2.1	68	1 MT3_MOUSE	P28184 mus musculus
8	5	1.8	10	1 HTF_HELZE	P16353 heliothis z
9	5	1.8	18	1 F1B8_ANAPL	P12802 anas platyr
10	5	1.8	28	1 V103_VACCP	Q00334 vaccinia vi
11	5	1.8	29	1 CX0D_CONMA	Q26350 conus magus
12	5	1.8	32	1 PSAM_MARPO	P31590 marchantia
13	5	1.8	37	1 CG2S_LUPIN	P09930 lupinus ang
14	5	1.8	39	1 ARG8_BOVIN	P80513 bos taurus
15	5	1.8	44	1 BAB4_DABBO	P14201 babesia bov
16	5	1.8	44	1 MF02_SCHPO	P34069 schizosacch
17	5	1.8	49	1 YD03_ARCFU	O28966 archaeoglob
18	5	1.8	51	1 DAPD_KLEPN	P41397 klebsicilla
19	5	1.8	51	1 YC33_OFNHO	Q9mtn9 oenothera h
20	5	1.8	52	1 TPCS_PRODO	P81074 prototenus
21	5	1.8	52	1 Y039_TREPA	O83081 treponema p
22	5	1.8	52	1 YABQ_ECOLI	P39221 escherichia
23	5	1.8	53	1 YF84_HAEIN	P44211 haemophilus
24	5	1.8	55	1 ATP8_STRUP	P15997 strongyloce
25	5	1.8	55	1 CX41_CBRAP	O46589 cebus apell
26	5	1.8	55	1 CX41_SAIUS	O46590 samirist
27	5	1.8	55	1 FER_CLOTH	P07508 clostridium
28	5	1.8	57	1 GNSB_ECOLI	P77695 escherichia
29	5	1.8	57	1 R132_ANASP	O8yqv6 anabaena sp
30	5	1.8	59	1 YQ08_RACSU	P54373 bacillus su
31	5	1.8	61	1 YA92_ARCFU	O29173 archaeoglob
32	5	1.8	62	1 YA95_ARCFU	O29170 archaeoglob
33	5	1.8	63	1 YA84_ARCFU	O29181 archaeoglob

34	5	1.8	64	1 Y314_ARCFU	O29931 archaeoglob
35	5	1.8	64	1 Y316_ARCFU	O29929 archaeoglob
36	5	1.8	64	1 YA74_ARCFU	O29189 archaeoglob
37	5	1.8	65	1 AMPL_MELGA	P80391 metagris g
38	5	1.8	66	1 YAEF_ECOLI	P52099 escherichia
39	5	1.8	67	1 SAS2_BACSU	P04832 bacillus su
40	5	1.8	67	1 YP0B_BACSU	P35151 bacillus su
41	5	1.8	68	1 MT3_HORSE	P37360 equus caball
42	5	1.8	68	1 MT3_HUMAN	P25713 homo sapien
43	5	1.8	69	1 APLC_RACFI	P22483 bacillus fi
44	5	1.8	69	1 CSRA_PSEAF	P95459 pseudomonas
45	5	1.8	70	1 SAS1_BACST	P06552 bacillus st
46	5	1.8	70	1 YXX8_PYRAB	O9v0b8 pyrococcus
47	4	1.4	10	1 RPP8_HGTIN	P30426 holthrops in
48	4	1.4	10	1 TRMK_RANIE	P56423 rana tempor
49	4	1.4	11	1 LPW_THETH	P05624 thermus the
50	4	1.4	11	1 PVK1_PERAM	P41837 periparata
51	4	1.4	11	1 RRP12_HAV	P13179 chardipura
52	4	1.4	14	1 GR75_CANFA	P99502 canis famli
53	4	1.4	14	1 PSAG_CUCSA	P42049 cucumis sat
54	4	1.4	15	1 C1QA_RAI	P31720 ratus norv
55	4	1.4	15	1 GR78_HORSE	P16392 equus cabal
56	4	1.4	15	1 ONC1_ONCMY	P83287 oncorhynch
57	4	1.4	16	1 ODO2_BOVIN	P11179 bos taurus
58	4	1.4	16	1 RL6_VIBPR	O56715 vibrio prot
59	4	1.4	16	1 TRYP_FELCA	P81071 felis silve
60	4	1.4	17	1 GAST_MACMU	P33714 macaca mala
61	4	1.4	18	1 RL23_HALCU	P05975 halobacteri
62	4	1.4	19	1 A122_HORSE	P81217 equus cabal
63	4	1.4	19	1 F1BA_BTSBO	P14441 bison bonas
64	4	1.4	19	1 F1BB_HORSE	P14471 equus cabal
65	4	1.4	19	1 F1BB_LANGL	P14473 lana glama
66	4	1.4	19	1 F1BB_P-G	P14477 sus scrofa
67	4	1.4	19	1 F1BB_TAPTE	P24539 tapirus ter
68	4	1.4	19	1 F1BB_VULVU	P24482 vulpus vulp
69	4	1.4	20	1 DER6_DERPT	P49277 dermatophag
70	4	1.4	20	1 F1BB_ELEMA	P24538 elephas max
71	4	1.4	20	1 PSBW_WHEAT	P55957 triticum ae
72	4	1.4	21	1 F1BB_ANTAM	P14455 antilocapra
73	4	1.4	21	1 TKNC_CARAU	P25421 carassius a
74	4	1.4	21	1 YFLA_METVO	P42017 metanococc
75	4	1.4	22	1 CH60_BOVIN	P10811 bos taurus
76	4	1.4	22	1 HGL2_FASHE	P40530 fasciola he
77	4	1.4	23	1 ACHS_ELEEL	P09692 electrophor
78	4	1.4	23	1 CR41_LITICE	P56242 litorea cae
79	4	1.4	23	1 CR42_LITICE	P56243 litorea cae
80	4	1.4	23	1 CR43_LITICE	P56244 litorea cae
81	4	1.4	23	1 VG22_BPT2	P21596 bacterioph
82	4	1.4	24	1 AMAA_BACTR	P37356 bacillus th
83	4	1.4	24	1 CH60_RELVI	P26317 heliothis v
84	4	1.4	24	1 DHAG_COMTE	P80705 commanusa t
85	4	1.4	24	1 HEMT_LINRE	P23543 lingula ree
86	4	1.4	24	1 LPA2_BACSU	P08497 bacillus su
87	4	1.4	25	1 ACP_ALCAU	P80917 alcaligenes
88	4	1.4	25	1 ANI3_MESAU	P81050 mesocricetu
89	4	1.4	25	1 CR11_LITSP	P56226 litorea ap
90	4	1.4	25	1 CR13_LITICE	P56228 litorea cae
91	4	1.4	25	1 CR15_LITICE	P56240 litorea cae
92	4	1.4	25	1 CR1A_LITSP	P52104 litorea spl
93	4	1.4	26	1 R220_BUCAP	P46246 buchiera ap
94	4	1.4	27	1 AL20_CARMA	P81823 carcius ma
95	4	1.4	27	1 CXDB_CONTE	P24159 conus loxii
96	4	1.4	27	1 L52_ADE07	P05663 human adeno
97	4	1.4	27	1 NTRC_VIBAL	P19905 vibrio atqi
98	4	1.4	28	1 ARYC_NOCGL	P80308 nocardia q
99	4	1.4	28	1 ETX2_BACCE	P80368 bacillus ce
100	4	1.4	28	1 VIP_SCYCA	P09685 scytoschinu
101	4	1.4	28	1 YA79_ARCFU	O29184 archaeoglob
102	4	1.4	29	1 ATPA_BRYMA	P26995 bryopsis ma
103	4	1.4	29	1 BREE_RANES	P40841 rana esculi
104	4	1.4	29	1 HRJ_BOIJA	P20416 bothrops ja
105	4	1.4	29	1 TAT_HV123	P12510 human immu-n
106	4	1.4	30	1 CY35_DESAC	P81079 desulfuromo

107 4 1.4 30 1 1TR1_GITLA
108 4 1.4 30 1 P2CO_AKISP
109 4 1.4 31 1 CXMA_CONMR
110 4 1.4 31 1 FIBB_CANFA
111 4 1.4 31 1 NAPA_HUMAN
112 4 1.4 31 1 PETI_POPDE
113 4 1.4 31 1 PETM_CYACA
114 4 1.4 31 1 Y3KD_BPCHP
115 4 1.4 32 1 GLBA_LAMSP
116 4 1.4 32 1 LPID_ECOLI
117 4 1.4 32 1 LPID_EDMTA
118 4 1.4 32 1 PHSS_DESHN
119 4 1.4 32 1 Y160_RPT4
120 4 1.4 33 1 ACT_DICVI
121 4 1.4 33 1 ATP7_SOLTU
122 4 1.4 33 1 BR2E_RANES
123 4 1.4 33 1 DIE3_PIG
124 4 1.4 33 1 GAST_DIDMA
125 4 1.4 33 1 HOKU_RHOOP
126 4 1.4 33 1 PSRT_MALZE
127 4 1.4 33 1 RL28_XENLA
128 4 1.4 33 1 RL4_HALCU
129 4 1.4 33 1 SC63_CANFA
130 4 1.4 33 1 VTLL_RAT
131 4 1.4 33 1 Y474_BORBU
132 4 1.4 33 1 YLCH_RPB82
133 4 1.4 33 1 YLCH_ECOLI
134 4 1.4 34 1 BR2C_RANES
135 4 1.4 34 1 GAST_CAPHI

ALIGNMENTS

RESULT 1
ID FIBH_CEREL STANDARD: PRT: 21 AA.
AC P14468;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Cervus elaphus (Red deer), and
OC Cervus elaphus nelsoni (American elk).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860, 9864;
RN [1]
RP SEQUENCE.
RC SPECIES=C.e.elaphus;
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
RN [2]
RP SEQUENCE.
RC SPECIES=C.e.nelsoni;
RA Moss G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro: IPR002181; Fibrinogen_C.
DR PROSITE: PS00514; FIBRIN_AC_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.

FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 6 6 SULFATION.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2558 MW; FCEE745D98931627 CRC64;
Query Match 2.1%; Score 6; DP 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 DEEED 224
DB 7 DEEED 12
RESULT 2
ID HS98_NEUCR STANDARD: PRT: 29 AA.
AC P31540;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Heat shock protein HSP98 (fragment).
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE.
RC STRAIN=74A;
RX MEDLINE=931112646; PubMed=1472534;
RA Vassiliev A.O., Plesofsky-Vig N., Hrabcik R.;
RT "Isolation, partial amino acid sequence, and cellular distribution of
RT heat-shock protein hsp98 from Neurospora crassa.";
RL Biochim. Biophys. Acta 1156:1-6(1992).
CC -1- SUBCELLULAR LOCATION: MORE CONCENTRATED IN POLYRIBOSOMES THAN
CC IN MONORIBOSOMES. AND PREFERENTIALLY LOCALIZED IN THE LAMC
CC SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY.
DR PIR: S28174; S28174.
DR InterPro: IPR001270; Chaprinin_clpa/S.
DR PROSITE: PS00870; CLPA1; 1.
DR PROSITE: PS00871; CLPB2; PARTIAL.
KW Chapterone; Heat shock; ATP-binding.
FT NON_TER 1 1
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 3194 MW; 51A1B058B8108106 CRC64;
Query Match 2.1%; Score 6; DP 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 185 TTAAEY 190
DB 22 TTAAEY 27
RESULT 3
ID YF07_HAEIN STANDARD: PRT: 38 AA.
AC P44229;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Very hypothetical protein H1507 in Mu-like prophage F100 region.
GN H1507.
OC Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;

```

RX MEDLINE-95350630; PubMed-7542800;
RA Fieischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.I., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Saudek D.M., Brandon R.C.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT *Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.;
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR FMRL; U32827; AAC23166.1;
DR TIGR; H11507;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 38 AA; 4123 MW; C32DE70D18D0BELL CRC64;
Query Match 2.1%; Score 6; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 SATKAE 175
DB 18 SATKAE 23
|||||
RESULT 4
ID AOPA_SHEEP STANDARD; PRT; 46 AA.
AC O18851;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Amino oxidase [flavin-containing] A (EC 1.4.3.4) (Monoamine oxidase)
DE (MAO-A) (Fragment).
GN MAOA.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98251499; PubMed-9589594;
RA Cambridge L.M., Lumsden J.M., Sadtighi M., Galloway S.M.;
RT *A Scat polymorphism at the ovine monoamine oxidase A locus (MAOA)*;
RL Anim. Genet. 28:457-457(1997).
CC !- FUNCTION: Catalyzes the oxidative deamination of biogenic and
CC xenobiotic amines and has important functions in the metabolism of
CC neuroactive and vasoactive amines in the central nervous system
CC and peripheral tissues. MAO-A preferentially oxidizes biogenic
CC amines such as 5-hydroxytryptamine (5-HT), norepinephrine and
CC epinephrine (By similarity).
CC !- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O -> O(2) -> RCHO + NH(3) +
CC H(2)O(2).
CC !- COFACTOR: FAD.
CC !- SUBUNIT: MONOMER, HOMO- OR HETERODIMER (CONTAINING TWO SUBUNITS OF
CC SIMILAR SIZE). EACH SUBUNIT CONTAINS A COVALENTLY BOUND FLAVIN.
CC ENZYMATICALLY ACTIVE AS MONOMER (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: Mitochondrial outer membrane.
CC !- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF023613; AAC16911.1;
DR HSSP; P27338; LQGS
KW Oxidoreductase; Flavoprotein; FAD; Transmembrane; Mitochondrion;
KW Catecholamine metabolism; Neurotransmitter degradation.
FT NON_TER 1 1
FT NON_TER 46 46
SQ SEQUENCE 46 AA; 5349 MW; FD06EBDEE44F8EBA CRC64;
Query Match 2.1%; Score 6; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 ELPSPR 113
DB 19 ELPSPR 24
|||||
RESULT 5
ID YF93_HAEIN STANDARD; PRT; 51 AA.
AC P44264;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein H11593.
GN H11593.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE-95350630; PubMed-7542800;
RA Fieischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.I., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Saudek D.M., Brandon R.C.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT *Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.;
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32833; AAC23244.1;
DR TIGR; H11593;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 51 AA; 6021 MW; 241DEAB9EF9AA3B2 CRC54;
Query Match 2.1%; Score 6; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 177 EKDGVK 182
|||||

```


Db 33 EKDCVK 38

RESULT 6

PAFP_PHYAM STANDARD; PRT; 65 AA.

AC PB1478: 082728; PRT; 65 AA.

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Anti-fungal protein 1 precursor (PAFP-S).

CN AFPS-1.

OS Phytolacca americana (Common pokeweed).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;

OC Caryophyllales; Caryophyllaceae; Phytolaccaceae; Phytolacca.

OX NCBI_TaxID=3527;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Seed;

RA Liu Y., Ren F., Xu C., Zhao J.;

RI "The sequence of a cDNA encoding anti-fungal protein in Phytolacca americana."

RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

RL [2]

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Seed;

RA Liu Y., Wu G., Zhao J.;

RI "Chromosomal sequence of a gene encoding anti-fungal protein in Phytolacca americana."

RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RL [3]

RN [3]

RP SEQUENCE OF 28-65.

RC TISSUE=Seed;

RA Feng S.;

RI Submitted (JUN-1998) to the SWISS-PROT data bank.

CC -1- FUNCTION: POSSIBLE ANTIFUNGAL ACTIVITY.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: FOUND ONLY IN SEEDS.

CC -1- SIMILARITY: BELONGS TO THE AMP FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF048745; AAC05129.1; -

DR EMBL: AF105062; AAD17942.1; -

KW Plant defense; Fungicide; Signal.

FT SIGNAL 1 27

FT CHAIN 28 65 ANTI-FUNGAL PROTEIN 1.

FT SEQUENCE 65 AA: 6804 MW: 0073DE3ABDC5H5C CHC64;

Query Match 2.1%; Score 6; DB 1; Length 65;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 LLSVIS 129

Db 18 LLSVIS 23

RESULT 7

MT3_MOUSE

ID MT3_MOUSE STANDARD; PRT; 68 AA.

AC P28184;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Metallothionein-III (MT-III) (Growth inhibitory factor) (GTF).

CN MT3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA MEDLINE-92335292; PubMed-1631128;

RA Palmer R.D., Findley S.D., Whitmore T.F., Darnam D.M.;

RT "MT-III, a brain-specific member of the metallothionein gene family."

RL Proc. Natl. Acad. Sci. U.S.A. 89:6333-6337(1992).

RN [2]

RP SEQUENCE FROM N.A.

RC MEDLINE-94314230; PubMed-8039715;

RA Naruse S., Igarashi S., Furuya T., Kobayashi H., Miyake T.,

RA Tsuji S.;

RT "Structures of the human and mouse growth inhibitory factor-encoding genes."

RL Gene 144:283-287(1994).

RN [3]

RP STRUCTURE BY NMR OF 32-68.

RA MEDLINE-2144630; PubMed-11560491;

RA Oz G., Zangger K., Armitage I.M.;

RT "Three-dimensional structure and dynamics of a brain specific growth inhibitory factor: metallothionein-3."

RL Biochemistry 40:11433-11441(2001).

CC -1- FUNCTION: INHIBITS SURVIVAL AND NEURITE FORMATION OF CORTICAL NEURONS IN VIRO.

CC -1- FUNCTION: BINDS HEAVY METALS. CONTAINS ZINC AND COPPER ATOMS AND ONLY A NEGLIGIBLE AMOUNT OF CADMIUM (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: BRAIN.

CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M93310; AAA39529.1; -

DR EMBL: S72046; AAR31397.1; -

DR PIR: A46034; A46034

DR PDB: 1J19; G3-OCT-01.

DR MGD: MGI:97173; Mt3.

DR InterPro: IPR003019; Metallothion.

DR InterPro: IPR000006; Metallothion_vert.

DR Pfam: PF0013; metalthio; 1.

DR PRINTS: PRO0860; MTVERTEBRATE

DR PROSITE: PS00203; METALLOTHIONEIN_VRT; 1.

DR Metal-binding: Metal-thiolate cluster; Zinc; Copper; Acetylation; 3D-structure.

FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).

FT DOMAIN 2 30 BETA.

FT DOMAIN 31 68 ALPHA.

FT METAL 6 8 CLUSTER B (BY SIMILARITY).

FT METAL 8 8 CLUSTER B (BY SIMILARITY).

FT METAL 14 14 CLUSTER B (BY SIMILARITY).

FT METAL 16 16 CLUSTER B (BY SIMILARITY).

FT METAL 20 20 CLUSTER B (BY SIMILARITY).

FT METAL 22 22 CLUSTER B (BY SIMILARITY).

FT METAL 25 25 CLUSTER B (BY SIMILARITY).

FT METAL 27 27 CLUSTER B (BY SIMILARITY).

FT METAL 30 30 CLUSTER B (BY SIMILARITY).

FT METAL 34 34 CLUSTER A (BY SIMILARITY).

FT METAL 35 35 CLUSTER A (BY SIMILARITY).

FT METAL 37 37 CLUSTER A (BY SIMILARITY).

FT METAL 38 38 CLUSTER A (BY SIMILARITY).

FT METAL 42 42 CLUSTER A (BY SIMILARITY).

FT METAL 45 45 CLUSTER A (BY SIMILARITY).

FT METAL 49 49 CLUSTER A (BY SIMILARITY).

```

FT METAL 51 CLUSTER A (BY SIMILARITY).
FT METAL 64 CLUSTER A (BY SIMILARITY).
FT METAL 66 CLUSTER A (BY SIMILARITY).
FT METAL 67 CLUSTER A (BY SIMILARITY).
SQ SEQUENCE 68 AA; 7009 MW; 791AF60E38FED3CA CRC64;

Query Match 2.1%; Score 6; DB 1; Length 68;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 KAAEK 178
DB 58 KAAEK 63
|||||

RESULT 8
HTE_HELZE
ID HTE_HELZE STANDARD; PRT; 10 AA.
AC P16353;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Hypertrehalosaemic hormone (He2-HRH).
OS Heliothis zea (Corn earworm) (Bollworm).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Diptysia; Noctuoidea; Noctuidae; Heliothinae; Helioverpa.
OX NCBI_TaxID=7113;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=88326324; PubMed=3415690;
RA Jaife H., Raina A.K., Riley C.T., Fraser B.A., Bird T.G.,
RA Tsenq C.M., Zhang Y.S., Hayes D.K.;
RT *Isolation and primary structure of a neuropeptide hormone from
RT Heliothis zea with hypertrehalosemic and adipokinetic activities.*;
RL Biochem. Biophys. Res. Commun. 155:344-350(1988).
CC -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYPH (TREHALOSE IS
CC THE MAJOR CARBOHYDRATE IN THE HEMOLYPH OF INSECTS).
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR: A31571; A31571.
DR InterPro: IPR002047; AKH.
DR PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1096 MW; 8E70367865A5B9D1 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 SGWGN 277
DB 6 SGWGN 10
|||||

RESULT 9
FIBR_ANAPL
ID FIBR_ANAPL STANDARD; PRT; 18 AA.
AC P12802;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Anas platyrhynchos (Domestic duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neoognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE.

```

```

RX MEDLINE=85168193; PubMed=3983613;
RA Min Y., Ping Z., Yaoshi Z.;
RI *Purification and primary structures of duck fibrinopeptides A and
RI B.*;
RL Sci. Sin., B, Chem. Biol. Agric. Med. Earth Sci. 28:31-35(1985).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COPACIOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR PIR: JPO102; JPO102.
DR InterPro: IPR002181; Fibrinogen_C.
DR PROSITE: PS00514; FIBRIN_AG_C DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.
FT PEPTIDE 1 18
FT MOD_RES 1 1 FIBRINOPEPTIDE B.
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID.
FT NON_TER 18 18 SULFATION.
SQ SEQUENCE 18 AA; 2028 MW; B0F15E7768F8A.F9 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 DDEDE 220
DB 7 DDEDE 11
|||||

RESULT 10
VIO3_VACCP
ID VIO3_VACCP STANDARD; PRT; 28 AA.
AC Q00334;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Protein I3 (Fragment).
GN I3L.
OS Vaccinia virus (Strain 1.1VP).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=31531;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91066899; PubMed=2250685;
RA Ryazankina O.I., Shchelkunov S.N., Muravlev A.I., Netsova N.A.,
RA Mikryukov N.N., Gutorov V.V., Nikulin A.E., Kulichkov V.A.,
RA Malygin E.G.;
RT *Molecular-biological study of vaccinia virus genome. II.
RT Localization and nucleotide sequence of vaccinia virus genes coding
RT for proteins 36K and 12K.*;
RL Mol. Biol. (Mosk) 24:968-976(1990).
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED IN THE EARLY AS WELL AS
CC THE LATE PHASE OF INFECTION.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X61165; CAA43473.1;
DR PIR: PS0395; PS0395.
KW Fatty protein; Late protein.
FT NON_TER 1 1
SQ SEQUENCE 28 AA; 3238 MW; C110813AC544F01C CRC64;

```

```

Query Match          1.8%; Score 5; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 DIDDE 218
DB 10 DIDDE 14

RESULT 11
CXOD_CONNA          STANDARD;          PRI: 29 AA.
AC Q26350;
DT 15-DEC-1998 (Rel. 37, Created)
DI 15-DEC-1998 (Rel. 37, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Omega-conotoxin MV1ID precursor (SNX-238) (Fragment).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6492;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150815; PubMed=8107968;
RA Monje V.D., Haack J.A., Naisbitt S.R., Miljanich G., Ramachandran J.,
RA Nasdasi L., Olivera B.M., Hillyard D.R., Gray W.R.;
RA "A new Conus peptide ligand for Ca channel subtypes.";
RI Neuropharmacology 32:1141-1149(1993);
RN [2]
RP STRUCTURE BY NMR.
RX PubMed=9920728;
RA Civera C., Vazquez A., Sevilla J.M., Bruix M., Gago F., Garcia A.G.,
RA Sevilla P.;
RA "Solution structure determination by two-dimensional 1H NMR of
RA Omega-conotoxin MV1ID, a calcium channel blocker peptide.";
RI Biochem. Biophys. Res. Commun. 254:32-35(1999).
CC -1- FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
CC and block voltage-sensitive calcium channels (VSCC). This toxin
CC blocks channels of the N-type as well as other types.
CC -1- SUBCELLULAR LOCATION: Secreted
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: S69322; AAR29902.1;
CC HSSP: P05484; LMV;
CC Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;
CC Amidation.
CC
CC NON_TER 1 1
CC FT PROPEP <1 3
CC FT PEPTIDE 4 28 OMEGA-CONOTOXIN MV1ID.
CC FT DISULFID 4 19
CC FT DISULFID 11 23
CC FT DISULFID 18 28
CC FT MOD_RES 28 28
CC SEQUENCE 29 AA; 3104 MW; 9E04B2EA1779CH22 CRC64;

Query Match          1.8%; Score 5; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.1e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GRGAS 259
DB 6 GRGAS 10

```

```

RESULT 12
PSAM_MARPO          STANDARD;          PRI: 32 AA.
AC P31590;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem I; reaction centre subunit X11 (PSI-M).
GN PSAM.
OS Marchantia polymorpha (Liverwort).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiaceae;
OC Marchantiaceae; Marchantia.
OX NCBI_TaxID=3197;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89068686; PubMed=2974085;
RA Umesono K., Inokuchi H., Shiki Y., Takeuchi M., Chang Z., Fukuawa H.,
RA Kohchi T., Shirai H., Ohyama K., Ozeki H.;
RA "Structure and organization of Marchantia polymorpha chloroplast
RA genome. II. Gene organization of the large single copy region from
RA rps12 to atpB.";
RI J. Mol. Biol. 203:299-331(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX Ohyama K., Fukuawa H., Kohchi T., Shirai H., Sano T., Sano S.,
RA Umesono K., Shiki Y., Takeuchi M., Chang Z., Inokuchi H.,
RA Ozeki H.;
RA "Chloroplast gene organization deduced from complete sequence of
RA Liverwort Marchantia polymorpha chloroplast DNA.";
RI Nature 322:572-574(1986).
RN [3]
RP -1- SIMILARITY: BELONGS TO THE PSAM FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X04465; CA82807.1;
CC PIR: A05015; A05015.
CC PIR: S0583; S0583.
CC Photosystem I; Photosynthesis; Chloroplast.
CC SEQUENCE 32 AA; 3565 MW; 674D9F557F5ADA5 CRC64;

Query Match          1.8%; Score 5; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.4e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ILLSV 127
DB 12 ILLSV 16

RESULT 13
CG2S_LUPAN          STANDARD;          PRI: 37 AA.
AC P09930;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Conglutin delta-2 small chain.
OS Lupinus angustifolius (Narrow-leaved blue lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OX NCBI_TaxID=3871;
RN [1]
RP SEQUENCE.

```

```

RC STRAIN-cv. White;
RA Lilley G.G., Inglis A.S.;
RT *Amino acid sequence of conglutinin delta, a sulfur-rich seed protein
RI of Lupinus angustifolius L. Sequence homology with the C-III alpha-
RT amylase inhibitor from wheat.*;
RL FEBS Lett. 195:235-241(1986).
CC -1- SUBUNIT: DIMER OF A SMALL CHAIN AND A LARGE CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
DR PIR: A23617; A23617.
KW Seed.
FT DISULFID 8 8 INTERCHAIN (WITH C-29 OF LARGE CHAIN).
FT DISULFID 20 20 INTERCHAIN (WITH C-17 OR C-18 OF LARGE
FT CHAIN).
FT DOMAIN 29 37 GLU/GLN-RICH.
FT SEQUENCE 37 AA: 4598 MW; 89784D55A5A1493A CRC64;

Query Match 1.8%; Score 5; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 EEEED 224
DB 33 EEEED 37

RESULT 14
ARGR_BOVIN
ID ARGR_BOVIN STANDARD; PRT: 39 AA.
AC P80513.
DT 01-OCT-1996 (Rel. 34, Created)
DI 01-OCT-1996 (Rel. 34, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Arginine-rich protein (Fragment).
GN ARMET OR ARP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=96239137; PubMed=860651;
RA Galat A., Gerbod M.C., Bouet F., Riviere S.;
RT *Proteins and their amino acid compositions: uniqueness, variability,
RI and applications.*;
RL Arch. Biochem. Biophys. 330:229-237(1996).
CC -1- SIMILARITY: BELONGS TO THE ARP FAMILY.
FT NON_TER 1 1
FT NON_TER 39 39
FT SEQUENCE 39 AA: 4589 MW; 00B3DE6206218E1C CRC64;

Query Match 1.8%; Score 5; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 TFSFA 140
DB 26 TFSFA 30

RESULT 15
BAB4_BABBO
ID BAB4_BABBO STANDARD; PRT: 44 AA.
AC P14201;
DT 01-JAN-1990 (Rel. 13, Created)
DI 01-JAN-1990 (Rel. 13, Last sequence update)
DI 01-APR-1990 (Rel. 14, Last annotation update)
DE High molecular weight antigen (Fragment).
OS Babesia bovis.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae; Babesia.
OX NCBI_TaxID=4865;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87201673; PubMed=3033495;
RA Gill A., Tams P., Kemp D.J.;
RT *cDNA clone encoding a high molecular weight antigen of Babesia
RI bovis.*;
RL Mol. Biochem. Parasitol. 22:195-202(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29838; AAA27795.1;
DR InterPro; IPR001813; 60s_ribosomal.
DR Pfam; PF00428; 60s_ribosomal; 1.
KW Antigen.
FT NON_TER 1 1
FT SEQUENCE 44 AA: 4731 MW; 3022E1E4AAFA0AF0 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 EEEED 224
DB 32 EEEED 36

RESULT 16
MFPM2_SCHPO
ID MFPM2_SCHPO STANDARD; PRT: 44 AA.
AC P34069;
DT 01-FEB-1994 (Rel. 28, Created)
DI 01-FEB-1994 (Rel. 28, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE M-factor precursor 2.
GN MFPM2 OR SPAC513.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.; AND SEQUENCE OF 33-41.
RX MEDLINE=92192030; PubMed=1547790;
RA Davey J.;
RT *Mating pheromones of the fission yeast Schizosaccharomycetes pombe:
RI purification and structural characterization of M-factor and
RL isolation and analysis of two genes encoding the pheromone.*;
RL FMOB J. 1:951-960(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., William R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayes J., Baker S., Hashar D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth J., Clancher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.T., Hunt S., Jakes K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean Z.,
RA Mooney P., Moule S., Murgall K., Murphy L., Niblett B., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares K., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grynopreuz B.,
RA Weljens I., Vanstreels E., Kieger M., Schaefer M., Mueller-Aber S.,
RA Gabel C., Fuchs M., Fritze C., Heizer E., Moestl D., Hilbert H.,

```

RA Horzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu A., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Baga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen O., Potashkin J.,
 RA Spakovski G.V., Uasey D., Barrell B.G., Nurse P.,
 RT "The genome sequence of *Schizosaccharomyces pombe*."
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: M-FACTOR IS A MATING PHEROMONE PRODUCED BY M-TYPE MATING
 CC CELLS. ALL THREE M-H GENES CONTRIBUTE TO THE PRODUCTION OF
 CC M-FACTOR.
 CC -1- INDUCTION: BY NITROGEN STARVATION. IT IS FURTHER INDUCED BY A
 CC PHEROMONE SIGNAL. ITS TRANSCRIPTION IS LIMITED TO M CELLS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X63628; CAA45176.1; .
 DR EMBL; X63628; CAA45176.1; .
 DR EMBL; AL122032; CAB58726.1; .
 DR PIR: S21028; S21028.
 DR InterPro: IPR001230; Prenyl_Site.
 KW Pheromone; Prenylation; Inoprotein; Multigene family; Methylation.
 FT PROPEP 1 32
 FT PEPTIDE 33 41 M-FACTOR.
 FT PROPEP 42 44 REMOVED IN MATURE FORM.
 FT LIPOD 41 41 FARNESYL
 FT MOD_RES 41 41 METHYLATION.
 SQ SEQUENCE 44 AA: 4858 MW: 7007642876877E CRC64:
 Query Match 1.8%; Score 5; DB 1; Length 44;
 Best Local Similarity 100.0%; Pred. No. 5.1e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 277 NTHSS 281
 Db 7 NTHSS 11
 RESULT 17
 YD03_ARCFU STANDARD; PRT; 49 AA.
 AC 028966;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1303.
 GN AF1303.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9369475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kraljic A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Iac N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reisch C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback I.,
 RA Cotton M.D., Spriggs T., Arlitch P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olson G.J., Fraser C.M., Smith H.O., Woese C.R.,

RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*."
 RL Nature 390:364-370(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE001014; AAB89952.1; .
 DR EMBL; AE001014; AAB89952.1; .
 DR TIGR: AEL303; .
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 49 AA: 6152 MW: FC9627F9HD3CB6D7 CRC64:
 Query Match 1.8%; Score 5; DB 1; Length 49;
 Best Local Similarity 100.0%; Pred. No. 5.1e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 130 LLNEP 134
 Db 25 LLNEP 29
 RESULT 18
 DAPD_KLEPN STANDARD; PRT; 51 AA.
 AC P41397;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE 2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase
 DE (EC 2.3.1.117) (Tetrahydropyridine N-succinyltransferase;
 DE (IHP succinyltransferase) (tetrahydropyridine succinylase)
 DE (Fragment).
 GN DAPD.
 OS Klebsiella pneumoniae.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Klebsiella.
 OX NCBI_TaxID=573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W5a1;
 RX MEDLINE=95272528; PubMed=7753028;
 RA Edwards R.A., Merrick M.J.;
 RT "The role of uridylyltransferase in the control of *Klebsiella*
 RT pneumoniae nif gene regulation."
 RL Mol. Gen. Genet. 247:169-198(1995).
 CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + 2,3,4,5-tetrahydropyridine-2-
 CC carboxylate -> CoA + N-succinyl-L-2-amino-6-oxohexanoate.
 CC -1- PATHWAY: FOURTH STEP IN THE BIOSYNTHESIS OF DIAMINOPIMELATE AND
 CC LYSINE FROM ASPARTATE SEMIALDEHYDE.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NOBL FAMILY OF
 CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G X(4).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X78685; CAA55354.1; .
 DR EMBL; X78685; CAA55354.1; .
 DR HSSP: P56220; 3DIT.
 DR InterPro: IPR001451; Hexapep_Transf.
 DR PROSITE: PS00101; HEXAPEP_TRANSFERASE; PARTIAL.
 KW Transferase; Acyltransferase; Repeat; Lysine biosynthesis;
 KW Diaminopimelate biosynthesis.

```

FT NON_TER 51
SQ SEQUENCE 51 AA: 5582 MW: 2F887878165FD5D CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 PANVD 143
    (||||)
DB 20 PANVD 24

RESULT 19
YCX3_OENHO
ID YCX3_OENHO STANDARD; PRI: 51 AA.
AC Q9MTN9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Hypothetical 5.4 kDa protein in ycf3 intronic region (ORF51).
OS Oenothera hookeri (Hooker's evening primrose).
OX Chloroplast.
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC eucosids II; Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=85636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Johansen;
RX MEDLINE=20309318; PubMed=10852478;
RA Hupfer H., Swiatek M., Hornung S., Hertzmann R.G., Maier R.M.,
RA Chiu W.-L., Sears B.;
RT *Complete nucleotide sequence of the Oenothera elata plastid
RT chromosome, representing plastome I of the five distinguishable
RT Eoenothera plastomes.*;
RI Mol. Gen. Genet. 263:581-585(2000).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ271079; CAB67136.1; -.
DR Chloroplast; Hypothetical protein.
KW SEQUENCE 51 AA: 5430 MW: DDD4F0749D1B25A1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 51;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 TILLS 126
    (||||)
DB 6 TILLS 10

RESULT 20
TPCS_PRODO
ID TPCS_PRODO STANDARD; PRI: 52 AA.
AC PB1074;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Troponin C, skeletal muscle (Fragment).
OS Protopterus dolloi (Lungfish).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Dipnoi; Lepidosireniformes; Protopteridae; Protopterus.
OX NCBI_TaxID=27779;
RN [1]
RP SEQUENCE.
RX MEDLINE=97442870; PubMed=9297801;

```

```

RA Francois J.M., Altintas A., Gerday C.;
RI *Characterization of the single tyrosine containing troponin C from
RI lungfish white muscle. Comparison with several fast skeletal muscle
RI troponin C's from fish species.*;
RL Comp. Biochem. Physiol. 117B:589-598(1997).
CC -I- FUNCTION: TROPONIN IS THE CENTRAL REGULATORY PROTEIN OF STRIATED
CC MUSCLE CONTRACTION. IT CONSISTS OF THREE COMPONENTS: IN-1 WHICH
CC IS THE INHIBITOR OF ACTOMYOSIN ATPASE, IN-2 WHICH CONTAINS THE
CC BINDING SITE FOR TROPOMYOSIN AND IN-3. THE BINDING OF CALCIUM TO
CC TN-C ABOLISHES THE INHIBITORY ACTION OF IN ON ACTIN FILAMENTS.
CC -I- MISCELLANEOUS: SKELETAL MUSCLE TROPONIN C BINDS FOUR CALCIUM IONS.
CC -I- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
DR HSSP; P02586; ITN4.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PSC0018; EF_HAND; 1.
KW Muscle protein; Calcium-binding; Repeat.
FT NON_TER 1
FT CA_BIND 15 26 EF-HAND 3.
FT NON_TER 52 52
SQ SEQUENCE 52 AA: 5949 MW: 6B8A8E88CA3F79CB CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 AEIIR 166
    (||||)
DB 28 AEIIR 32

RESULT 21
Y039_TREPA
ID Y039_TREPA STANDARD; PRI: 52 AA.
AC Q83081;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0039.
GN TP0039.
OS Treponema pallidum.
CC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Dodson R., Norris S.J., Weinstock G.M., White G., Sutton G.C.,
RA Fraser C.M., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Otterback T.,
RA McDonald L., Artlich P., Bowman C., Colton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman C., Smith H.C.,
RA Venter J.C.;
RT *Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.*;
RI Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF001189; AAC65040.1; -.
DR TIGR; TP0039; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 52 AA: 5934 MW: 6C4C8666A4E85E1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 NVRTI 123
DB 31 NVRTI 35
|||||

RESULT 22
YABQ_ECOLI
ID YABQ_ECOLI STANDARD; PRT; 52 AA.
AC P39221;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yabQ.
GN YABQ OR B0057.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision: Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN (1)
RN SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-92334977; PubMed-1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Tshihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region."
KL Nucleic Acids Res. 20:3305-3308(1992).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Godden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RN IDENTIFICATION.
RA Rudd K.F.;
RL Unpublished observations (NOV-1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR PWWL: D10483; NOT_ANNOTATED_CDS.
DR EMBL: AK000116; AAC73168.1;
DR EcoGene: K012611; yabQ.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 52 AA; 5737 MW; 8079697D5AC6E4A5 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 LLNEP 134
DB 47 LLNEP 51
|||||

RESULT 23
YB84_HAEIN
ID YB84_HAEIN STANDARD; PRT; 53 AA.
AC P44211;
DT 01-NOV-1995 (Rel. 32, Created)

```

```

CI-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Very hypothetical protein H1484 in Mu-like prophage FlaMu region.
GN H1484.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision: Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN (1)
RN SEQUENCE FROM N.A.
RC STRAIN-Rd / KW23 / ATCC 51907;
RX MEDLINE-95350630; PubMed-7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-P., Dougherty H.A., Merrick C.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Lita L.I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Sprague D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Goehagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Seall K.V., Fraser C.M., Smith H.C.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32826; AAC23339.1;
DR TIGR: H11484;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 53 AA; 6359 MW; BD868885A9D5D4L4 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 DEEPF 223
DB 49 DEEPF 53
|||||

RESULT 24
ATP8_STRPU
ID ATP8_STRPU STANDARD; PRT; 55 AA.
AC P5937;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A61).
GN MTATP8 OR ATP8.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Eucchinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN (1)
RN SEQUENCE FROM N.A.
RX MEDLINE-8901951; PubMed-3172215;
RA Jacobs H.T., Elliott D.J., Math V.B., Farquharson A.;
RT "Nucleotide sequence and gene organization of sea urchin
RL J. Mol. Biol. 202:185-227(1988).
CC -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) -> ADP + phosphate +
CC H(+) (Out).

```

CC -!- SUBCELLULAR LOCATION: Membrane-bound.
 CC -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC FMH; X12631; CAA31156.2; -
 DR PIR: S01504; S01504.
 DR InterPro: IPR001421; ATPase8_mit.
 DR Pfam: PF00895; ATP-synt_8; 1.
 KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
 SQ SEQUENCE 55 AA: 6252 MW: 84DA0DD4FF543B6F CRC64;

 CC Query Match 1.8%; Score 5; DB 1; length 55;
 CC Best Local Similarity 100.0%; Pred. No. 5.6e-02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 127 VISIL 131
 DB 23 VISIL 27

 CC RESULT 25
 CC CX41_CERAP STANDARD: PRT: 55 AA.
 AC 046589;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome c oxidase subunit IV isoform 1 (EC 1.9.3.1) (COX IV-1)
 DE (Cytochrome c oxidase polypeptide IV) (Fragment).
 GN COX41 OR COX4.
 OS Cebus apella (brown-capped capuchin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebus.
 OX NCBI_TaxID=9515;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97277139; PubMed-9115172;
 RA Wu W., Goodman M., Lomax M.I., Grossman L.I.;
 RT "Molecular evolution of cytochrome c oxidase subunit IV: evidence for
 RT positive selection in simian primates.";
 RL J. Mol. Evol. 44:477-491(1997).
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) -> 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE IV FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF042778; AAB97758.1; -
 DR InterPro: IPR004203; COX4.
 DR Pfam: PF02936; COX4; 1.
 KW Oxidoreductase; Inner membrane; Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 55 55
 SQ SEQUENCE 55 AA: 6277 MW: 7231A7FE4D81A773 CRC64;

 CC Query Match 1.8%; Score 5; DB 1; length 55;
 CC Best Local Similarity 100.0%; Pred. No. 5.6e-02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 127 VISIL 131
 DB 23 VISIL 27

CC Best Local Similarity 100.0%; Pred. No. 5.6e-02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 SOKAL 13
 DB 33 SOKAL 37

 CC RESULT 26
 CC CX41_SAUS STANDARD: PRT: 55 AA.
 AC 046590;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome c oxidase subunit IV isoform 1 (EC 1.9.3.1) (COX IV-1)
 DE (Cytochrome c oxidase polypeptide IV) (Fragment).
 GN COX41 OR COX4.
 OS Saimiri ustus (Squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
 OX NCBI_TaxID=66265;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97277139; PubMed-9115172;
 RA Wu W., Goodman M., Lomax M.I., Grossman L.I.;
 RT "Molecular evolution of cytochrome c oxidase subunit IV: evidence for
 RT positive selection in simian primates.";
 RL J. Mol. Evol. 44:477-491(1997).
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) -> 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE IV FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF042779; AAB97759.1; -
 DR InterPro: IPR004203; COX4.
 DR Pfam: PF02936; COX4; 1.
 KW Oxidoreductase; inner membrane; Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 55 55
 SQ SEQUENCE 55 AA: 6320 MW: AF91A7FE4D99DE94 CRC64;

 CC Query Match 1.8%; Score 5; DB 1; length 55;
 CC Best Local Similarity 100.0%; Pred. No. 5.6e-02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 SOKAL 13
 DB 33 SOKAL 37

 CC RESULT 27
 CC FER_CLOTH STANDARD: PRT: 55 AA.
 AC P07508;
 DT 01-APR-1988 (Rel. 37, Created)
 DT 01-APR-1988 (Rel. 37, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Ferredoxin.
 OS Clostridium thermocellum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.

NCBI_TaxID=1515;
SEQUENCE.
RA Bruschi M., Cambillau C., Bovier-Lapierre G.E., Bonicel J.J.,
RA Forget P.;
*Sequence determination and three-dimensional modelling of
RI Clostridium thermocellum ferredoxin: structural considerations for its
RL high thermal stability.*;
RL Blochm. Biophys. Acta 873:31-37(1986).
CC -1- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.
DR PIR: A24932; A24932.
DR USSP: P00198; 2FDN.
DR InterPro: IPR001450; 4Fe4S_ferredoxin.
DR Pfam: PF00037; fer4.2.
DR PROSITE: P000353; 4FE4SFEROXIN.
KW Electron transport; iron-sulfur; Repeat: 4Fe-4S.
FT METAL 8 8 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 11 11 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 14 14 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 18 18 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 37 37 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 40 40 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 43 43 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 47 47 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 55 AA: 6306585468584213 CRC64;
Query Match 1-8%; Score 5; DH 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 265 APQOK 269
|||||
DB 51 APQOK 55

RESULT 28
GNSB_ECOLI
ID GNSB_ECOLI STANDARD; PRI: 57 AA.
AC P77695;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GnsB protein.
CN GNSB OR B1550.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
The complete genome sequence of Escherichia coli K-12.;
RT Science 277:1453-1474(1997).
RI [2]
RN STRAIN-K12;
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sanpei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
*A 570-kb DNA sequence of the Escherichia coli K-12 genome

corresponding to the 28.0-40.1 min region on the linkage map.*;
RL DNA Res. 3:363-377(1996).
RN [3]
RP FUNCTION.
RX MEDLINE=2429224; PubMed=11544213;
RA Sugai K., Shimizu H., Nishiyama K., Ito, Ikada H.;
*Overexpression of yccL (gnsA) and ydfY (gnsB) increases levels of
RT unsaturated fatty acids and suppresses both the temperature-sensitive
RT fabA6 mutation and cold-sensitive secG null mutation of Escherichia
RT coli.*;
RL J. Bacteriol. 183:5523-5528(2001).
CC -1- FUNCTION: Overexpression increases levels of unsaturated fatty
CC acids and suppresses both the temperature sensitive fabA6 mutation
CC and cold-sensitive secG null mutation.
CC -1- SIMILARITY: BELONGS TO THE GNS FAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF000252; AAC74623.1; ALT INIT.
CC ENBL; D90798; BAA15253.1; ALT INIT.
CC EcoGene; EGI3837; gnsB.
CC Complete proteome.
KW SEQUENCE 57 AA: 6548 MW: 1750BD19C3B9186A CRC64;
Query Match 1-8%; Score 5; DH 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 TKAE 176
|||||
DB 8 TKAE 12

RESULT 29
RL32_ANASP
ID RL32_ANASP STANDARD; PRI: 57 AA.
AC O8YQY6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L32.
CN RPF OR RPL32 OR ASL3674.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21595285; PubMed=11759840;
RX Watanabe A., Nakamura Y., Wolk C.P., Kuritz J., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakasaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
*Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.*;
RI DNA Res. 8:205-213(2001).
RN [2]
CC -1- SIMILARITY: BELONGS TO THE L32P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AP003593; BAB75373.1;

DR InterPro: IPR002677; Ribosomal_L32p.
 DR Pfam: PF01783; Ribosomal_L32p; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 57 AA; 6502 MW; 8A96103B9F22C210 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 57;
 Best Local Similarity 100.0%; Pred. No. 5.8e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEF 222
 D 11111
 Db 53 EDEEF 57

RESULT 30
 YQDR_BACSU
 ID YQDR_BACSU STANDARD; PRT; 59 AA.
 AC P54373;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yqdh.
 GN YQDR.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=95219086; PubMed=7704261;
 RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.;
 RT "Complete nucleotide sequence of a skin element excised by DNA
 rearrangement during sporulation in *Bacillus subtilis*.";
 RL Microbiology 141:323-327(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=97124195; PubMed=8969508;
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
 RA Kobayashi Y.;
 RT "Systematic sequencing of the 283 kb 210 degrees-212 degrees region of
 the *Bacillus subtilis* genome containing the skin element and many
 sporulation genes.";
 RL Microbiology 142:3103-3111(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunat F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Horviss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Hrouillet S., Bruchsi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gham S.Y., Glaser P., Goffeau A., Golightly P.J., Grand G.,
 RA Guisepi G., Guy R.J., Haga K., Haicich J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones I.,
 RA Joris R., Karamata D., Kashara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinou S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Ito H., Masuda S., Maue C., Medigic C.,
 RA Medina N., Mellado R.P., Mizuno M., Moostl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shit B.S., Soldo B.,
 RA Sorokin A., Tacconi K., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Welleneger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus
 subtilis*.";
 RL Nature 390:249-256(1997).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
 or send an email to license@isb-sib.ch)

EMBL: D32216; NOT_ANNOTATED_CDS.
 DR EMBL: D84432; CAB12409.1;
 DR EMBL: Z99117; CAB14546.1;
 DR Subtilist; BG11512; yqdh.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 59 AA; 6739 MW; 8A7EF81AC97779 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 VISLL 13;
 D 11111
 Db 24 VISLL 28

RESULT 3;
 YA92_ARCFU
 ID YA92_ARCFU STANDARD; PRT; 61 AA.
 AC Q29273;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AFI092.
 GN AFI092.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Hickey E.K., Peterson J.D.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerevage A.R., Graham D.E., Kyriakos N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty H.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil I.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback I.,
 RA Cotton M.D., Sprague J., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997)
 CC -!- SIMILARITY: BELONGS TO THE UPF0165 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
 or send an email to license@isb-sib.ch)

EMBL: AE001028; AB090155.1;

```

DR TIGR: AF1092;
DR InterPro: IPR002807; DUF104;
DR Pfam: PF01954; DUF104; 1;
DR ProDom: PD005964; DUF104; 1;
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 61 AA; 6920 MW; ABB2ABA9645CB03C CRC64;

Query Match 1.8%; Score 5; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 6.2e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 IYENG 90
DB 8 IYENG 12

RESULT 32
YA95_ARCFU STANDARD; PRT; 62 AA.
AC 029170;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF1095.
GN AF1095.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervagge A.R., Graham D.F., Kyrides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeck R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: BELONGS TO THE UPF0165 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AF001028; AAB90154.1;
CC TIGR: AF1095;
CC InterPro: IPR002807; DUF104.
CC Pfam: PF01954; DUF104; 1.
CC ProDom: PD005964; DUF104; 1.
CC Hypothetical protein: Complete proteome.
CC SEQUENCE 62 AA; 7435 MW; 9562AC3C7D69C6C7 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 62;
Best Local Similarity 100.0%; Pred. No. 6.3e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 IYENG 90
DB 8 IYENG 12

```

```

RESULT 33
YA84_ARCFU STANDARD; PRT; 63 AA.
AC 029181;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF1084.
GN AF1084.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervagge A.R., Graham D.F., Kyrides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeck R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: BELONGS TO THE UPF0165 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AE001028; AAB90156.1;
CC TIGR: AF1084;
CC InterPro: IPR002807; DUF104.
CC Pfam: PF01954; DUF104; 1.
CC ProDom: PD005964; DUF104; 1.
CC Hypothetical protein: Complete proteome.
CC SEQUENCE 63 AA; 7619 MW; D7957FD026134AA; CRC64;

Query Match 1.8%; Score 5; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 6.4e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 IYENG 90
DB 8 IYENG 12

RESULT 34
Y314_ARCFU STANDARD; PRT; 64 AA.
AC 029931;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0314.
GN AF0314.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE-98049343; PubMed-9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White G., Nelson K.F.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus S.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: BELONGS TO THE UPF0165 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001083; AAB90918.1; ALT_INIT.
DR TIGR: AF0316;
DR InterPro: IPR002807; DUF104.
DR Pfam: PF01954; DUF104; 1.
DR ProDom: PD005964; DUF104; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 64 AA: 7255 MW: 4405986D9618D4A6 CRC64;
Query Match 1.8%: Score 5; DB 1; Length 64;
Best Local Similarity 100.0%: Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 IYENG 90
DB 8 IYENG 12
|||||
RESULT 35
Y316_ARCFU STANDARD; PRI: 64 AA.
AC Q2929;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0316.
GN AF0316.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID-2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE-98049343; PubMed-9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White G., Nelson K.F.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Kyripides N.C.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus S.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: BELONGS TO THE UPF0165 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001083; AAB90923.1;
DR TIGR: AF0314;
DR InterPro: IPR002807; DUF104.
DR Pfam: PF01954; DUF104; 1.
DR ProDom: PD005964; DUF104; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 64 AA: 7255 MW: 4405986D9618D4A6 CRC64;
Query Match 1.8%: Score 5; DB 1; Length 64;
Best Local Similarity 100.0%: Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 IYENG 90
DB 8 IYENG 12
|||||

```

```

RT reducing archaeon Archaeoglobus fulgidus.
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: BELONGS TO THE UPF0165 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001083; AAB90918.1; ALT_INIT.
DR TIGR: AF0316;
DR InterPro: IPR002807; DUF104.
DR Pfam: PF01954; DUF104; 1.
DR ProDom: PD005964; DUF104; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 64 AA: 7512 MW: F5D64D542561D797 CRC64;
Query Match 1.8%: Score 5; DB 1; Length 64;
Best Local Similarity 100.0%: Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 IYENG 90
DB 8 IYENG 12
|||||
RESULT 36
YA74_ARCFU STANDARD; PRI: 64 AA.
AC Q29189;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF1074.
GN AF1074.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID-2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE-98049343; PubMed-9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White G., Nelson K.F.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Kyripides N.C.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus S.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: BELONGS TO THE UPF0165 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF001029; AAB90182.1;
DR TIGR: AF1074;
DR InterPro: IPR002807; DUF104.

```

```

DR Pfan; PF01954; DUF104; 1.
DR ProDom; PD005964; DUF104; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 64 AA; 7448 MW; 6H7E8D26DFCDEFD4C CRC64;

Query Match 1.8%; Score 5; DB 1; Length 64;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 IYENG 90
DB 8 IYENG 12

RESULT 37
ID AMPL1_MELGA STANDARD; PRI: 55 AA.
AC P80391;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Antimicrobial peptide Thp1 precursor (Turkey heterophil peptide 1).
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=bone marrow.
RX MEDLINE=98418188; PubMed=9745566;
RA Brockus C.W., Harmon B.G., Jackwood M.W.;
RT "Characterization of beta-defensin prepropeptide mRNA from chicken
and turkey bone marrow.";
RL Anim. Genet. 29:283-289(1998).
RN [2]
RP SEQUENCE OF 26-60.
RX MEDLINE=95053386; PubMed=7964174;
RA Evans E.W., Beach G.G., Wunderlich J., Harmon B.G.;
RT "Isolation of antimicrobial peptides from avian heterophils.";
RL J. Leukoc. Biol. 56:661-665(1994).
CC - FUNCTION: BACTERICIDAL ACTIVITY; INHIBITS STAPHYLOCOCCUS AUREUS
AND ESCHERICHIA COLI.
CC - SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF033337; AAC36053.1;
CC Antibiolic; Signal. 25
FT SIGNAL 1
FT CHAIN 26 60 ANTIMICROBIAL PEPTIDE THP1.
FT PROPEP 61 65
FT DISULFID 31 53 BY SIMILARITY.
FT DISULFID 38 59 BY SIMILARITY.
FT DISULFID 43 60 BY SIMILARITY.
SQ SEQUENCE 65 AA; 7086 MW; 9237FC0F7B448D45 CRC64;

Query Match 1.8%; Score 5; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 LSVIS 129
DB 46 LSVIS 50

RESULT 38
ID YAEF_ECOLI STANDARD; PRI: 57 AA.

```

```

ID YAEF_ECOLI STANDARD; PRI: 66 AA.
AC P52099;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yaeP.
GN YAEF OR B0189.1.
GS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Yamamoto Y.;
AL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=96839317; PubMed=9723924;
RA Pichoff S., Aliand L., Guedant A., Castanie M.-P., Rouché J.-P.;
RT "An Escherichia coli gene (yaeP) suppresses temperature-sensitive
mutations in essential genes by modulating rho-dependent
transcription termination.";
RI Mol. Microbiol. 29:859-869(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatiner F.R., Plunkett G. III, Bloch C.A., Perna N.L., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew S.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RI Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Takamoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada F., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Sangei G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
4.0 - 6.0 min (189,987 - 281,416bp) region.";
RI Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; D49445; BAA08430.1;
CC EMBL; Z50870; CAA90753.1;
CC EMBL; AE000128; NOT_ANNOTATED_CDS.
CC EMBL; D83536; BAA77865.1;
CC ECoGene; HG33222; yaeP.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 66 AA; 7214 MW; 98AC59C7178E2D9F CRC64;

Query Match 1.8%; Score 5; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 258 ASGDL 262
DB 16 ASGDL 20

RESULT 39
ID SAS2_BACSU STANDARD; PRI: 57 AA.

```

AC 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Small, acid-soluble spore protein B (SASP).
 GN SSPB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 ON NCBI_TaxID=1423;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96195826; PubMed=3009398;
 RA Connors M.J., Mason J.M., Setlow P.;
 RT "Cloning and nucleotide sequencing of genes for three small, acid-
 soluble proteins from *Bacillus subtilis* spores."
 RL J. Bacteriol. 166:417-425(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boursier R., Boursier L., Brans A., Braun M., Brignol S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter K.M.,
 RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,
 RA Medina N., Mellado R.P., Mizuno M., Mosti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pulic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Koche B., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin H.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*."
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: SASP ARE FOUND TO SPORE DNA. THEY ARE DOUBLE-STRANDED
 CC DNA-BINDING PROTEINS THAT CAUSE DNA TO CHANGE TO AN A-LIKE
 CC CONFORMATION. THEY PROTECT THE DNA BACKBONE FROM CHEMICAL AND
 CC ENZYMATIC CLEAVAGE AND ARE THUS INVOLVED IN DORMANT SPORE'S HIGH
 CC RESISTANCE TO UV LIGHT.
 CC -!- MISCELLANEOUS: SASP ARE DEGRADED IN THE FIRST MINUTES OF SPORE
 CC GERMINATION AND PROVIDE AMINO ACIDS FOR BOTH NEW PROTEIN SYNTHESIS
 CC AND METABOLISM.
 CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-TYPE SASP FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>)
 CC or send an email to license@ebi.ac.uk.

CC EMBL: M12621; AAA22834.1;
 DR EMBL: Y14080; CAA74447.1;
 DR EMBL: Z99109; CAG12814.1;
 DR PIR: H24546; H24546.
 DR Subtilisin; BGI0787; ssfB.
 DR InterPro; IPR001448; SASP.
 DR Pfam; PF00269; SASP_1.
 DR PROSITE; PS00304; SASP_1; 1.
 DR PROSITE; PS00684; SASP_2; 1.
 KW DNA-binding; Sporulation; Multigene family; Complete proteome.
 FT SITE 25 26 CLEAVAGE (BY SPORE PROTEASE).
 SQ SEQUENCE 67 AA: 6980 MW: 19A3972301E81621 CRC64;
 Query Match 1.88; Score 5; DH 1; length: 67;
 Best Local Similarity 100.08; Pred. No. 6,7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 AQQQM 6
 DB 58 AQQQM 62
 ID YPUB_BACSU STANDARD; PRT; 67 AA.
 AC P35151;
 DT 01-FEB-1994 (Rel. 28, Created)
 DI 01-FEB-1994 (Rel. 28, Last sequence update)
 DI 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ypuB (OREX1).
 GN YPUB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / Marburg;
 RX MEDLINE=95020538; PubMed=7934829;
 RA Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;
 RT "The organization of the *Bacillus subtilis* 169 chromosome region
 between the *spoVA* and *serA* genetic loci, based on sequence data."
 RL Mol. Microbiol. 10:385-395(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boursier R., Boursier L., Brans A., Braun M., Brignol S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter K.M.,
 RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,
 RA Medina N., Mellado R.P., Mizuno M., Mosti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pulic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Koche B., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin H.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*."
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: SASP ARE FOUND TO SPORE DNA. THEY ARE DOUBLE-STRANDED
 CC DNA-BINDING PROTEINS THAT CAUSE DNA TO CHANGE TO AN A-LIKE
 CC CONFORMATION. THEY PROTECT THE DNA BACKBONE FROM CHEMICAL AND
 CC ENZYMATIC CLEAVAGE AND ARE THUS INVOLVED IN DORMANT SPORE'S HIGH
 CC RESISTANCE TO UV LIGHT.
 CC -!- MISCELLANEOUS: SASP ARE DEGRADED IN THE FIRST MINUTES OF SPORE
 CC GERMINATION AND PROVIDE AMINO ACIDS FOR BOTH NEW PROTEIN SYNTHESIS
 CC AND METABOLISM.
 CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-TYPE SASP FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>)
 CC or send an email to license@ebi.ac.uk.

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L09228; AAA67476.1; -;
 DR EMBL; Z99116; CAB14266.1; -;
 DR PIR; S45538; S45538.
 DR Subtilist; BGI0513; ypuB.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 67 AA; 7237 MW; 1C0DBC56627EE5FE CRC64;

Query Match 1.8%; Score 5; DB 1; Length 67;
 Best local Similarity 100.0%; Pred. No. 6.7e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LMLEL 17
 |
 |
 |
 |
 |
 DB 20 LMLEL 24

Search completed: April 10, 2003, 10:39:49
 Job time : 12.2988 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein protein search, using sw model

Run on: April 10, 2003, 10:37:10 ; Search time 20.9012 seconds

(without alignments)

1297.052 Million cell updates/sec

Title: US-09-930-026-2

Perfect score: 282

Sequence: 1 MAQQQMTSSQKALMLELKL.....AMAPQPKPIIISGWGTHSSC 282

Scoring table: OLICC

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 20812

Minimum DB seq length: 0

Maximum DB seq length: 70

Post-processing: Listing first 135 summaries

Database:

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	2.8	59	2	S56692
2	7	2.5	21	4	A32902
3	6	2.1	26	2	F37396
4	6	2.1	26	2	E37396
5	6	2.1	29	2	S28174
6	6	2.1	31	2	T70361
7	6	2.1	38	2	H64033
8	6	2.1	39	2	PQ0011
9	6	2.1	41	2	T14350
10	6	2.1	43	2	S02031
11	6	2.1	45	2	S35286
12	6	2.1	49	2	A56388
13	6	2.1	50	2	F95032
14	6	2.1	51	2	F66796
15	6	2.1	51	2	D64037
16	6	2.1	56	2	T50373
17	6	2.1	68	2	A46534
18	6	2.1	68	2	I67866
19	5	1.8	10	2	A35571
20	5	1.8	15	2	A31902
21	5	1.8	16	2	S13898
22	5	1.8	18	2	JP0102
23	5	1.8	22	2	G83924
24	5	1.8	23	2	T44539
25	5	1.8	24	2	T04142
26	5	1.8	25	2	A60807
27	5	1.8	25	2	S27229
28	5	1.8	25	2	A61499
29	5	1.8	26	2	S11629

30	5	1.8	26	2	A33413	cytochrome f - tur
31	5	1.8	29	2	A58537	omega-conotoxin MV
32	5	1.8	30	2	I57689	ubiquinol-cytochro
33	5	1.8	31	2	T44925	hypothetical prote
34	5	1.8	32	2	A05015	hypothetical prote
35	5	1.8	34	2	F70242	hypothetical prote
36	5	1.8	34	2	S12554	hydroxymethylgluta
37	5	1.8	37	2	A23617	T-kininogenase (EC
38	5	1.8	37	2	A23617	conglutin delta-2
39	5	1.8	38	2	C82482	hypothetical prote
40	5	1.8	38	2	A44862	microtubule associ
41	5	1.8	38	2	T46593	phytoene dehydroge
42	5	1.8	39	2	E42799	photosystem I chai
43	5	1.8	39	2	S69268	arginine-rich prot
44	5	1.8	40	2	S44136	NADH2 dehydrogenas
45	5	1.8	42	2	S77815	probable DNA lias
46	5	1.8	42	2	T07285	hypothetical prote
47	5	1.8	42	2	C86487	hypothetical prote
48	5	1.8	43	2	I50608	hypothetical prote
49	5	1.8	43	2	B97505	hypothetical prote
50	5	1.8	44	2	A84184	hypothetical prote
51	5	1.8	44	2	S21028	m-factor precursor
52	5	1.8	44	2	A54524	high molecular wei
53	5	1.8	44	2	F97940	conserved hypotet
54	5	1.8	45	2	T07433	hypothetical prote
55	5	1.8	45	2	T52272	R2R3-MYB transcrip
56	5	1.8	48	2	R05249	collagen alpha 2(I
57	5	1.8	48	2	D64614	hypothetical prote
58	5	1.8	48	2	E71901	hypothetical prote
59	5	1.8	49	2	F69412	hypothetical prote
60	5	1.8	51	2	S54757	2,3,4,5-tetrahydro
61	5	1.8	51	2	E95992	hypothetical prote
62	5	1.8	51	2	R81443	highly acidic prot
63	5	1.8	51	2	A62564	hypothetical prote
64	5	1.8	52	2	T43384	hypothetical prote
65	5	1.8	52	2	F71373	hypothetical prote
66	5	1.8	52	2	A64727	yabQ protein - Esc
67	5	1.8	53	2	E64031	hypothetical prote
68	5	1.8	53	2	D83832	hypothetical prote
69	5	1.8	53	2	AH2383	hypothetical prote
70	5	1.8	54	2	B97870	hypothetical prote
71	5	1.8	55	2	A24932	terredoxin 2[4Fe-4
72	5	1.8	55	2	S01504	H-transporting tw
73	5	1.8	55	2	A70070	hypothetical prote
74	5	1.8	56	2	PQ0437	genome polyprotein
75	5	1.8	56	2	PQ0436	genome polyprotein
76	5	1.8	56	2	D86540	hypothetical prote
77	5	1.8	56	2	E72081	hypothetical prote
78	5	1.8	56	2	A59424	attractin - Aplysi
79	5	1.8	56	2	AB2573	hypothetical prote
80	5	1.8	57	2	AC2265	50S ribosomal prot
81	5	1.8	58	2	D33356	prothymosin alpha
82	5	1.8	58	2	F33356	prothymosin alpha
83	5	1.8	58	2	A64910	hypothetical prote
84	5	1.8	59	2	F33356	prothymosin alpha
85	5	1.8	59	2	D89008	prothymosin alpha
86	5	1.8	59	2	B70243	protein W08A12.2 l
87	5	1.8	59	2	E69950	hypothetical prote
88	5	1.8	59	2	G71145	hypothetical prote
89	5	1.8	59	2	F98276	hypothetical prote
90	5	1.8	60	2	C36852	A388 protein - var
91	5	1.8	60	2	D72168	A39R protein - var
92	5	1.8	60	2	AF1832	hypothetical prote
93	5	1.8	61	2	C69386	conserved hypotet
94	5	1.8	61	2	A43430	alpha IIB beta 3 l
95	5	1.8	62	2	F69386	conserved hypotet
96	5	1.8	62	2	A26878	tyrocidine synthet
97	5	1.8	62	2	A69034	hypothetical prote
98	5	1.8	62	2	H97839	hypothetical prote
99	5	1.8	62	2	A02151	hypothetical prote
100	5	1.8	63	2	C69385	conserved hypotet
101	5	1.8	63	2	S09643	probable cytochrom
102	5	1.8	63	2	B45574	liber - canine ade

103 5 1.8 63 2 069115
 104 5 1.8 64 2 065289
 105 5 1.8 64 2 065384
 106 5 1.8 64 2 112828
 107 5 1.8 64 2 097915
 108 5 1.8 65 2 038601
 109 5 1.8 65 2 075236
 110 5 1.8 65 2 136041
 111 5 1.8 65 2 083700
 112 5 1.8 66 2 089810
 113 5 1.8 66 2 083266
 114 5 1.8 67 2 069718
 115 5 1.8 67 2 090236
 116 5 1.8 67 2 084870
 117 5 1.8 67 2 066426
 118 5 1.8 67 2 045538
 119 5 1.8 67 2 001333
 120 5 1.8 67 2 089987
 121 5 1.8 67 2 082097
 122 5 1.8 68 2 046034
 123 5 1.8 68 2 044392
 124 5 1.8 68 2 070002
 125 5 1.8 68 2 087516
 126 5 1.8 69 2 095333
 127 5 1.8 69 2 017721
 128 5 1.8 69 2 083236
 129 5 1.8 69 2 086894
 130 5 1.8 69 2 069385
 131 5 1.8 70 2 025234
 132 5 1.8 70 2 056784
 133 5 1.8 70 2 090857
 134 5 1.8 70 2 090856
 135 5 1.8 70 2 090855

ALIGNMENTS

RESULT 1
 S56692
 hypothetical protein (clone AFN2) - wild cat (fragment.)
 C:Species: *Avina fatua* (wild cat)
 C:Date: 10-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 21-Jun-2000
 C:Accession: S56692
 R:Johnson, R.R.; Cravens, H.J.; Chaverra, M.E.; Dyer, W.E.
 Plant Mol. Biol. 28, 113-122, 1995
 A:Title: Characterization of cDNA clones for differentially expressed genes in embryos of
 A:Reference number: S56690; MUID:95306780; PMID:7787175
 A:Accession: S56692
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-59 <J01>
 A:Cross-references: EMBL:U19999; NID:q726475; PIDN:AAA6741.1; PID:q726476
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

Query Match 2.8%; Score 8; DH 2; Length 59;
 Best Local Similarity 100.0%; Pred. No. 0; Gaps 0;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 DDDEEE 223
 DB 43 DDDEEE 50

RESULT 2
 A32902
 GMP reductase / glucose-6 phosphate 1-dehydrogenase fusion protein - human (fragment)
 C:Species: *Homo sapiens* (man)
 C:Date: 10-Sep-1998 #sequence_revision 10-Sep-1998 #text_change 30-Jun-2001
 C:Accession: A32902
 R:Kanno, H.; Huang, I.Y.; Kan, Y.W.; Yoshida, A.
 Cell 58, 595-606, 1989
 A:Title: two structural genes on different chromosomes are required for encoding the map

A:Reference number: A32902; MUID:89336791; PMID:2758458
 A:Accession: A32902
 A:Molecule type: protein
 A:Residues: 1-21 <KAN>
 R:Yoshida, A.; Kan, Y.W.
 Cell 62, 11-12, 1990
 A:Title: Origin of "fused" glucose-6-phosphate dehydrogenase.
 A:Reference number: A55177; MUID:90304899; PMID:1694726
 A:Contents: annotation
 C:Comment: This peptide was sequenced from a fragment of a chimeric protein produced
 was later shown to be an artifact of purification and not a natural component of er
 C:Keywords: fusion protein
 F:1-13/Region: GMP reductase
 F:14-21/Region: glucose-6-phosphate 1-dehydrogenase

Query Match 2.5%; Score 7; DH 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 GASGDLA 263
 DB 14 GASGDLA 20

RESULT 3
 F37396
 pollen allergen Poa p I - Kentucky bluegrass (fragment.)
 C:Species: *Poa pratensis* (Kentucky bluegrass)
 C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
 C:Accession: F37396
 R:Esch, R.E.; Klapper, D.G.
 Mol. Immunol. 26, 557-561, 1989
 A:Title: Isolation and characterization of a major cross-reactive grass group I alle
 A:Reference number: A37396; MUID:89364850; PMID:2475768
 A:Accession: F37396
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-26 <ESC>
 C:Superfamily: expansin
 C:Keywords: pollen

Query Match 2.1%; Score 6; DH 2; Length 26;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 TKAEAE 177
 DB 7 TKAEAE 12

RESULT 4
 E37396
 pollen allergen Agr a I - bent grass (fragment)
 C:Species: *Agrostis alba* (bent grass)
 C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jun-2000
 C:Accession: E37396
 R:Esch, R.E.; Klapper, D.G.
 Mol. Immunol. 26, 557-561, 1989
 A:Title: Isolation and characterization of a major cross-reactive grass group I alle
 A:Reference number: A37396; MUID:89364850; PMID:2475768
 A:Accession: E37396
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-26 <ESC>
 C:Superfamily: expansin
 C:Keywords: pollen

Query Match 2.1%; Score 6; DH 2; Length 26;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 TKAEAE 177
 DB 7 TKAEAE 12

Db 7 IKAEE 12

RESULT 5

heat-shock protein hsp98 - Neurospora crassa (fragment)
 N:Alternate names: ATP-dependent G-P proteinase regulatory chain
 N:Contains: adenosinetriphosphatase (EC 3.6.1.3)
 C:Species: Neurospora crassa
 C>Date: 22-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 30-Apr-1999
 C:Accession: S28174
 R:Vassiliev, A.G.; Plesofsky-Vig, N.; Hrambl, R.
 Biochim. Biophys. Acta 1156, 1-6, 1992
 A:Title: Isolation, partial amino acid sequence, and cellular distribution of heat-shock
 A:Reference number: S28174; MUID:93112646; PMID:1472534
 A:Accession: S28174
 A:Molecule type: protein
 A:Residues: 1-29 <VAS>
 C:Function:
 C:Description: allows clip to hydrolyze polypeptides and proteins, probably by a chaperon
 e activity; ATP hydrolysis is required for Clip hydrolysis of proteins but not of smaller
 C:Superfamily: endopeptidase G-P ATP-binding chain
 C:Keywords: ATP; heat shock; hydrolyase; nucleotide binding; stress-induced protein

Query Match

Best Local Similarity 2.1% Score 6; DB 2; Length 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 TLAAY 190

|||||

Db 22 TLAAY 27

RESULT 6

enkephalin precursor - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jan-2000
 C:Accession: IT0381
 R:Watkinson, A.; Dockray, G.J.; Young, J.; Gregory, H.
 Biochim. Biophys. Acta 955, 231-235, 1988
 A:Title: Characterisation of N-terminally extended met-enkephalin 6-Arg, 7-Gly, 8-Leu va
 A:Reference number: J0381; MUID:88281515; PMID:3395626
 A:Accession: J0381
 A:Molecule type: protein
 A:Residues: 1-31 <WAT>
 C:Superfamily: proenkephalin
 C:Keywords: neuropeptide; opioid peptide

Query Match

Best Local Similarity 2.1% Score 6; DB 2; Length 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 201 NSSLL 206

|||||

Db 10 NSSLL 15

RESULT 7

hypothetical protein H1507 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
 C:Accession: E64013
 R:Flaeschmann, R.D.; Mans, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.; Glodok, A.; Kelley, J.M.; Weidman, J.
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Guehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.G.; Venter,
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: E64013
 A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-38 <TIGR>

A:Cross-references: GB:032827; GB:142023; NID:q1574343; PIDN:AAC23166.1; PID:q157435

C:Genetics:

A:Start codon: GTG

Query Match

Best Local Similarity 2.1% Score 6; DB 2; Length 38;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 170 SATKAE 175

|||||

Db 18 SATKAE 23

RESULT 8

tubulin beta chain - carrot (fragment)
 C:Species: Daucus carota (carrot)
 C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 12-Apr-1995
 C:Accession: PQ0011
 R:Okamura, S.; Azumano, I.
 Biochem. Int. 16, 1103-1109, 1988
 A:Title: Primary structure of the carboxy-terminal region of a higher plant beta tub
 A:Reference number: PQ0011; MUID:89025927; PMID:3178860
 A:Accession: PQ0011
 A:Molecule type: mRNA
 A:Residues: 1-39 <OKA>
 A:Experimental source: strain GD2
 C:Comment: the carboxy-terminal region of tubulin is highly acidic.
 C:Superfamily: tubulin

Query Match

Best Local Similarity 2.1% Score 6; DB 2; Length 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 218 EDPEEE 223

|||||

Db 29 EDPEEE 34

RESULT 9

tubulin beta-3 chain - carrot (fragment)
 C:Species: Daucus carota (carrot)
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000
 C:Accession: T14350
 R:Okamura, S.; Sonohata, K.; Naito, K.; Ohkawa, H.; Kuramori, S.; Tatsuta, M.; Minar
 submitted to the EMBL Data Library, July 1996
 A:Description: Characterization of beta-tubulin genes of carrot.
 A:Reference number: 217999
 A:Accession: T14350
 A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-41 <OKA>

A:Cross-references: FMBL:064430; NID:q1553124; PID:q1553125

A:Experimental source: cultivar Kintoki

C:Comment: Tubulin is a dimer of alpha and beta chains and is found in the microtubule

changeable site on its beta chain and at a nonexchangeable site not yet identified.

C:Complex: heterodimer: alpha and beta chain

C:Superfamily: tubulin

C:Keywords: heterodimer

Query Match

Best Local Similarity 2.1% Score 6; DB 2; Length 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 218 EDPEEE 223

|||||

Db 32 EDPEEE 37

RESULT 10

S02031
 zinc-binding protein: - rat (tentative sequence) (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 31-Mar-2000
 C:Accession: S02031
 R:Brand, L.A.; Heinicke, A.; Kratzin, H.; Soelling, H.D.
 Eur. J. Biochem. 177, 561-568, 1988
 A:Title: Properties of a 19-kDa Zn(2+)-binding protein and sequence of the Zn(2+)-binding
 A:Reference number: S02031; MUID:890644809; PMID:3197718
 A:Accession: S02031
 A:Molecule type: protein
 A:Residues: 1-43 <R>
 C:Superfamily: prothymosin: alpha
 C:Keywords: zinc

Query Match 2.1% Score 6; DB 2; Length 43;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 218 EDEEEE 223
 DB 30 EDEEEE 35

RESULT 11
 S35286
 hypothetical protein 45 (eaE 3' region) - phage P22
 C:Species: phage P22
 C:Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 08-Oct-1999
 C:Accession: S35286
 R:Wulff, D.L.; Ho, Y.S.; Powers, S.; Rosenberg, M.
 Mol. Microbiol. 9, 261-271, 1993
 A:Title: The int genes of bacteriophages P22 and lambda are regulated by different mecha
 A:Reference number: S35286; MUID:94018622; PMID:8412679
 A:Accession: S35286
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-45 <W>
 A:Cross-references: EMBL:L06206; NID:g215287; PIDN:AA018886.1; PID:g215294

Query Match 2.1% Score 6; DB 2; Length 45;
 Best Local Similarity 100.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 55 GGYFRA 60
 DB 20 GGYFRA 25

RESULT 12
 A56388
 dopamine- and cAMP-regulated neuronal phosphoprotein DARPP-32 - rat (fragments)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
 C:Accession: A56388
 R:Dusdoulis, F.; Cohen, D.; Nairn, A.C.; Greengard, P.; Girault, J.A.
 J. Biol. Chem. 270, 8772-8778, 1995
 A:Title: Phosphorylation of DARPP-32, a dopamine- and cAMP-regulated phosphoprotein, by
 A:Reference number: A56388; MUID:95238371; PMID:7721783
 A:Accession: A56388
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-49 <DES>
 C:Keywords: phosphoprotein
 F:7.35.43/Binding site: phosphate (Ser) (covalent) (by casein kinase I) #status experime

Query Match 2.1% Score 6; DB 2; Length 49;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 217 DEDEEE 222
 DB 24 DEDEEE 29

RESULT 13
 F95032
 hypothetical protein SP0277 [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: F95032
 R:Tettelein, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.;
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
 on, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, H.A.; Morri
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: F95032
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-50 <KOR>
 A:Cross-references: GB:AF005672; PIDN:AAK74455.1; PID:g14971749; GSPDB:GN00164; TIGR
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP0277

Query Match 2.1% Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 121 RTILLS 126
 DB 25 RTILLS 30

RESULT 14
 I66796
 fetal troponin T 3 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 13-Aug-1999
 C:Accession: I66796
 R:Briggs, M.M.; Schachet, F.
 Dev. Biol. 158, 503-509, 1993
 A:Title: Origin of fetal troponin T: Developmentally regulated splicing of a new exc
 A:Reference number: I53073; MUID:93345743; PMID:8344466
 A:Accession: I66796
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-51 <RES>
 A:Cross-references: EMBL:U04980; NID:g440820; PIDN:AAA16033.1; PID:g440821
 C:Superfamily: troponin T

Query Match 2.1% Score 6; DB 2; Length 51;
 Best Local Similarity 100.0%; Pred. No. 1e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 218 EDDEEE 223
 DB 37 EDDEEE 42

RESULT 15
 D64037
 hypothetical protein H11593 - Haemophilus influenzae (strain H0 KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
 C:Accession: D64037
 R:Feischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavag
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidma
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vent
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: D64037

A:Status: nucleic acid sequence not shown; translation not shown.

A:Molecule type: DNA

A:Residues: 1-51 <100>

A:CROSS-references: GB:952833; GR:142023; NID:gl574432; PIDN:AAC23244.1; PID:gl574442;

Query Match 2.1%; Score 6; DB 2; Length 31;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 177 EKQGVK 182

DB 33 EKQGVK 38

RESULT 16

153573

fetal troponin T 2 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 13-Aug-1999

C:Accession: F53073

R:Briggs, M.M.; Schachet, F.

A:Title: Origin of fetal troponin 1: Developmentally regulated splicing of a new exon in

A:Reference number: F53073; MUID:93345743; PMID:8344466

A:Accession: F53073

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-56 <85>

A:CROSS-references: EMBL:J04979; NID:q440818; PIDN:AAA16032.1; PID:q440819

C:Superfamily: Troponin T

Query Match 2.1%; Score 6; DB 2; Length 56;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223

DB 42 EDEEE 47

RESULT 17

A46034

reticulothionein 3, brain-specific - mouse

R:Alternate names: neurotrophic growth inhibitory factor

C:Species: Mus musculus (house mouse)

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999

C:Accession: A46034

R:Palmiter, R.D.; Findley, S.D.; Whitmore, T.E.; Durnam, D.M.

Proc. Natl. Acad. Sci. U.S.A. 89, 5333-5337, 1992

A:Title: MT-211, a brain-specific member of the metallothionein gene family.

A:Reference number: A46034; MUID:92335292; PMID:1631128

A:Accession: A46034

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-68 <PA5>

A:CROSS-references: GB:M93110; NID:gl94133; PIDN:AAA39529.1; PID:gl99134

A:Note: sequence extracted from NCBI backbone (NCBIN:108715, NCBIN:111115, NCBIN:108716)

C:Superfamily: metallothionein

Query Match 2.1%; Score 6; DB 2; Length 68;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 KAAEAK 178

DB 58 KAAEAK 63

RESULT 18

167866

growth inhibitory factor - mouse

C:Species: Mus sp. (mouse)

C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 20-Aug-1999

C:Accession: I67866

R:Naruse, S.; Igarashi, S.; Furuya, T.; Kobayashi, H.; Miyatake, T.; Tsuji, S.

Gene 144, 283-287, 1994

A:Title: Structures of the human and mouse growth inhibitory factor-encoding genes.

A:Reference number: 153863; MUID:9414230; PMID:8039715

A:Accession: I67866

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-68 <RFS>

A:CROSS-references: GR:S72046; NID:q565191; PIDN:AAB31357.1; PID:q565192

C:Genetics:

A:Gene: GIF

A:Introns: 11/1; 33/1

C:Superfamily: metallothionein

Query Match 2.1%; Score 6; DB 2; Length 58;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 KAAEAK 178

DB 58 KAAEAK 63

RESULT 19

A31571

hypertrehalosemic/adipokinetic hormone - bollworm

N:Alternate names: Hez-WrTH

C:Species: Heliothis zea (bollworm, corn earworm, tomato fruitworm)

C:Date: 30-Jun-1989 #sequence_revision 23-Mar-1995 #text_change 31-Oct-1997

C:Accession: A31571

R:Jaife, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Bird, T.G.; Tserng, C.M.; Zhang,

Biochem. Biophys. Res. Commun. 155, 344-350, 1988

A:Title: Isolation and primary structure of a neuropeptide hormone from Heliothis ze

A:Reference number: A31571; MUID:88326324; PMID:3415690

A:Accession: A31571

A:Molecule type: protein

A:Residues: 1-10 <JAF>

C:Superfamily: adipokinetic hormone

C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglut

! F10/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

! F10/Modified site: amidated carboxyl end (Asn) #status experimental

Query Match 1.8%; Score 5; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 SGWGN 277

DB 6 SGWGN 10

RESULT 20

A31902

bone acidic glycoprotein-75 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 31-Dec-1993

C:Accession: A31902

R:Gorski, J.P.; Shimizu, K.

J. Biol. Chem. 263, 15938-15945, 1988

A:Title: Isolation of new phosphorylated glycoprotein from mineralized phase of bone

A:Reference number: A31902; MUID:89034045; PMID:2846530

A:Accession: A31902

A:Molecule type: protein

A:Residues: 1-15 <GOR>

A:Note: 14-Glu and 15-Glu were also found

C:Keywords: glycoprotein

Query Match 1.8%; Score 5; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 EDEED 224

```

Db      10 DEED 14
|||||
RESULT 21
S13898
C:Species: oryctolagus cuniculus (domestic rabbit)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
C:Accession: S13898
R:Fujimori-Arai, Y.; Koyama, I.; Hirano, K.; Sakagishi, Y.; Komoda, T.
Arch. Biochem. Biophys. 284, 320-325, 1991
A:Title: Purification and partial characterization of intestinal-like alkaline phosphatase
A:Reference number: S13898; MUID:91112827; PMID:1389515
A:Accession: S13898
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-16 <FUJ>
C:Keywords: phosphoric monoester hydrolase

Query Match      1.8%; Score 5; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 DEED 224
|||||
Db      4 DEED 8

RESULT 22
JP0102
fibrinogen beta chain - duck (fragment)
N:Contains: fibrinopeptide B
C:Species: Anas platyrhynchos (domestic duck)
C:Date: 30-Jun-1987 #sequence_revision 28-Dec-1987 #text_change 13-Sep-1996
C:Accession: JP0102
R:Min, Y.; Ping, Z.; Yaoshi, Z.
Sci. Sin. B Chem. Biol. Agric. Med. Earth Sci. 28, 31-35, 1985
A:Title: Purification and primary structures of duck fibrinopeptides A and B.
A:Reference number: A94238
A:Accession: JP0102
A:Molecule type: protein
A:Residues: 1-18 <MIN>
C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide
C:Keywords: blood coagulation; plasma; pyroglutamic acid; sulfoprotein
E1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
E2/Binding site: sulfate (Tyr) (covalent) #status predicted

Query Match      1.8%; Score 5; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 DDEDE 220
|||||
Db      7 DDEDE 11

RESULT 23
G83924
hypothetical protein RH2199 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G83924
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, E.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G83924
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <STO>
A:Cross-references: GB:AP001514; GB:BA000004; NID:gl0174613; PIDN:BAH05918.1; GSPDH:GN00
A:Experimental source: strain C-125

```

```

C:Genetics:
A:Gene: BH2199

Query Match      1.8%; Score 5; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 IFSPA 140
|||||
Db      2 IFSPA 6

RESULT 24
T44539
hypothetical protein V'R2 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 31-Jan-2000
C:Accession: T44539
R:Nakayama, K.; Takashima, K.; Ishihara, H.; Shinomiya, T.; Kageyama, M.; Kanaya, S.
submitted to the EMBL Data Library, August 1999
A:Description: Genetic relationship between bacteriophages and bacteriophages.
A:Reference number: 222790
A:Accession: T44539
A:Status: preliminary; translated from GH/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-23 <NAK>
A:Cross-references: EMBL:AB030825; PIDN:BAH83154.1
A:Experimental source: strain PA01
C:Genetics:
A:Note: V'R2

Query Match      1.8%; Score 5; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 276 GNTHS 280
|||||
Db      2 GNTHS 6

RESULT 25
T04142
tubulin beta-2 chain - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 13-Aug-1999
C:Accession: T04142
R:Okamura, S.; Okahara, K.; Iida, T.; Tanaka, M.; Nishimura, M.
submitted to the EMBL Data Library, February 1997
A:Description: Differential oscillation of the transcript levels of beta-tubulin isot;
A:Reference number: Z15146
A:Accession: T04142
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-24 <OKA>
A:Cross-references: EMBL:U91563; NID:gl907396; PIDN:AAB50566.1; PID:gl907397
A:Experimental source: strain Bright Yellow 2
C:Superfamily: tubulin

Query Match      1.8%; Score 5; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 222
|||||
Db      16 EDEEE 20

RESULT 26
A60807
heat shock protein 90 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C:Accession: A60807

```

R: Denis, M.
 Anal. Biochem. 173, 405-411, 1988
 A: Title: Two-step purification and N-terminal amino acid sequence analysis of the rat M-130 protein.
 A: Reference number: A60807; MUID: 89048319; PMID: 3189818
 A: Accession: A60807
 A: Molecule type: protein
 A: Residues: 1-25 <DEN>
 C: Comment: This protein associates with steroid hormone receptors and with the Rous sarcoma virus.
 C: Superfamily: heat shock protein 90
 C: Keywords: phosphoprotein

Query Match 1.8%; Score 5; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 EEPVE 26
 ||||
 DB 8 EEPVE 12

RESULT 27
 S27229
 prothymosin alpha homolog - Escherichia coli (fragments)
 C: Species: Escherichia coli
 C: Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
 C: Accession: S27229
 R: Vartapetian, A.; Chichkova, N.; Lyakhov, I.; Makarova, T.; Evstafieva, A.; Bogdanov, A.
 FEBS Lett. 313, 95-97, 1992
 A: Title: Segments of Escherichia coli genome similar to the exons of human prothymosin alpha.
 A: Reference number: S27229; MUID: 93050251; PMID: 1426289
 A: Accession: S27229
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-6; 7-25 <VAK>
 A: Note: the authors did not translate the codon for residues 17

Query Match 1.8%; Score 5; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 DEEE 223
 ||||
 DB 20 DEEE 24

RESULT 28
 A61499
 glutathione transferase (EC 2.5.1.18) - nematode (Haemonchus contortus) (fragment)
 C: Species: Haemonchus contortus
 C: Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-May-1999
 C: Accession: A61499
 R: Sharp, P.J.; Smith, D.R.J.; Bach, W.; Wagland, B.M.; Cobon, G.S.
 Int. J. Parasitol. 21, 839-846, 1991
 A: Title: Purified glutathione S-transferases from parasites as candidate protective antigens.
 A: Reference number: A61499; MUID: 92129048; PMID: 1774119
 A: Accession: A61499
 A: Status: preliminary
 A: Molecule type: protein
 A: Residues: 1-25 <SHA>
 C: Keywords: transferase

Query Match 1.8%; Score 5; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 AFIIR 166
 ||||
 DB 14 AFIIR 18

RESULT 29
 S11629
 elastin precursor - human (fragment)

C: Species: Homo sapiens (man)
 C: Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Feb-1997
 C: Accession: S11629
 R: Bressan, G.M.
 submitted to the EMBL Data Library, June 1989
 A: Reference number: S11629
 A: Accession: S11629
 A: Molecule type: DNA
 A: Residues: 1-26 <BRE>
 A: Cross-references: EMBL: X15603
 C: Comment: This may not be a genuine elastin sequence.
 C: Genetics:
 C: Map position: 2q31-qter
 C: Superfamily: elastin

Query Match 1.8%; Score 5; DB 2; Length 26;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 SILHP 99
 ||||
 DB 18 SILHP 22

RESULT 30
 A33413
 cytochrome f - turnip chloroplast (fragments)
 C: Species: chloroplast Brassica rapa (turnip)
 C: Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 20-Apr-2000
 C: Accession: A33413
 R: Morand, L.Z.; Frame, M.K.; Colvert, K.K.; Johnson, D.A.; Krogmann, D.W.; Davis, D.
 Biochemistry 28, 8039-8047, 1989
 A: Title: Plastocyanin cytochrome f interaction.
 A: Reference number: A33413; MUID: 90105304; PMID: 2605172
 A: Accession: A33413
 A: Status: preliminary
 A: Molecule type: protein
 A: Residues: 1-26 <MOR>
 C: Genetics:
 A: Genome: chloroplast
 C: Keywords: chloroplast

Query Match 1.8%; Score 5; DB 2; Length 26;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 ITLVD 34
 ||||
 DB 20 ITLVD 24

RESULT 31
 A58537
 omega-conotoxin MV1D precursor - cone shell (Conus magus) (fragment)
 C: Species: Conus magus (magus cone)
 C: Date: 27-Mar-1997 #sequence_revision 11-Apr-1997 #text_change 16-Jul-1999
 C: Accession: A58537
 R: Monje, V.D.; Haack, J.A.; Naisbitt, S.R.; Miljanich, G.; Ramachandran, J.; Nardas
 Neuropharmacology 32, 1141-1149, 1993
 A: Title: A new Conus peptide ligand for Ca channel subtypes.
 A: Reference number: A58537; MUID: 94150815; PMID: 8107968
 A: Accession: A58537
 A: Molecule type: mRNA
 A: Residues: 1-29 <MON>
 A: Cross-references: GB: S69322; NID: g545399; PIDN: AAB29902.1; PID: g545400
 A: Note: the predicted peptide was chemically synthesized and alternative disulfide
 C: Superfamily: omega-conotoxin
 C: Keywords: toxin; venom
 F: 4-29/Product: omega-conotoxin MV1D #status predicted <MAT>
 F: 4-19,11-23,18-28/Disulfide bonds: #status predicted

Query Match 1.8%; Score 5; DB 2; Length 29;
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GRCAS 259
|||||

Db 6 GRCAS 10

RESULT 32
157689
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - northern hairy-nosed wombat
C:Species: mitochondrion Lasiorhinus krefftii (northern hairy-nosed wombat).
C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 03-Jun-2002
C:Accession: I57689
R:Taylor, A.C.; Sherwin, W.B.; Wayne, R.K.
Mol. Ecol. 3, 277-290, 1994
A:Title: Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat.
A:Reference number: 157689; PMID:95005563; PMID:7921355
A:Accession: I57689
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-40 <TAY>
A:Cross-References: EMBL:U04964; NID:g441111; PIDN:AAA53176.1; PID:g441112
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; mitochondrion; oxidative phosph

Query Match 1.8%; Score 5; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 FSPAN 14;
|||||

Db 26 FSPAN 30

RESULT 33
144925
hypothetical protein nir [imported] - Anabaena sp. (strain PCC7120) (fragment)
C:Species: Anabaena sp.
A:Variety: strain PCC7120
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
C:Accession: T44925
R:Frías, J.E.; Flores, K.; Herrero, A.
J. Bacteriol. 179, 477-486, 1997
A:Title: Nitrate assimilation gene cluster from the heterocyst-forming Cyanobacterium Anabaena sp. strain PCC7120.
A:Reference number: 222870; PMID:97144534; PMID:8990301
A:Accession: T44925
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-31 <FRL>
A:Cross-References: EMBL:X59708; NID:g1483204; PIDN:CAA68047.1; PID:g1483205
A:Experimental source: strain PCC7120
C:Genetics:
A:Gene: nir
C:Superfamily: ferredoxin-nitrite reductase

Query Match 1.8%; Score 5; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 ARKX 180
|||||

Db 20 ARKX 24

RESULT 34
A05015
hypothetical protein 32 - liverwort (Marchantia polymorpha) chloroplast
C:Species: chloroplast Marchantia polymorpha
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 21-Jul-2000
C:Accession: S01583; A05015

R:Umesono, K.; Inokuchi, H.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Fukuzawa, H.; Kohen, J. Mol. Biol. 203, 299-331, 1988
A:Title: Structure and organization of Marchantia polymorpha chloroplast genome. II.
A:Reference number: S01567; MUID:89068686; PMID:2974085
A:Accession: S01583
A:Molecule type: DNA
A:Residues: 1-32 <UME>
A:Cross-References: EMBL:X04465; NID:g11640; PIDN:CAA28071.1; PID:g11658
R:Ohyama, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Nature 322, 572-574, 1986
A:Title: Chloroplast gene organization deduced from complete sequence of liverwort M. polymorpha.
A:Reference number: A38014
A:Contents: annotation: gene organization, sites, features
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 1.8%; Score 5; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ILLSV 127
|||||

Db 12 ILLSV 16

RESULT 35
F70242
hypothetical protein BB123 - Lyme disease spirochete plasmid I/lp26-4
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: F70242
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W. son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V. ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, R. Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70242
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-34 <KLE>
A:Cross-References: GB:AE000789; NID:g2690079; PIDN:AA66206.1; PID:g2690103; TIGR.R
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid

Query Match 1.8%; Score 5; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 LSVIS 129
|||||

Db 25 LSVIS 29

RESULT 36
S12554
hydroxymethylglutaryl-CoA reductase (NADPH2) (EC 1.1.1.34) - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C:Accession: S12554
R:Clarke, P.R.; Hardie, D.G.
EMBO J. 9, 2439-2446, 1990
A:Title: Regulation of HMG-CoA reductase: identification of the site phosphorylated in the active form.
A:Reference number: S12554; MUID:90316098; PMID:2369897
A:Accession: S12554
A:Molecule type: protein
A:Residues: 1-12; 13-34 <CLA>
C:Keywords: coenzyme A; NADP; oxidoreductase; phosphoprotein

Query Match 1.8%; Score 5; DB 2; Length 34;

```

Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 108 ELPSE 112
Db 26 ELPSE 30

RESULT 37
H35545
T-kininogenase (EC 3.4.21.-) heavy chain - rat (fragment)
N:Alternate names: proteinase B heavy chain
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 21-Mar-1996
C:Accession: H35545; E41429
R:Xiong, W.; Chen, L.M.; Chao, J.
C: Biol. Chem. 265, 2822-2827, 1990
A:Title: Purification and characterization of a kallikrein-like T-kininogenase.
A:Reference number: H35545; MUID:90153911; PMID:2303430
A:Accession: B35545
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-37 <X10>
R:Kato, H.; Nakanishi, E.; Eniyeji, K.; Hayashi, I.; Oh-Ishi, S.; Iwanaga, S.
J. Biochem. 102, 1389-1404, 1987
A:Title: Characterization of serine proteinases isolated from rat submaxillary gland: with special reference to trypsin and chymotrypsin
A:Reference number: A41429; MUID:88198057; PMID:3482210
A:Accession: E41429
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-18, 25, 21 <XAT>
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase

Query Match 1.8%; Score 5; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 179 DGKVV 183
Db 21 DGKVV 25

RESULT 38
A23617
conglutinin delta-2 small chain - narrow-leaved blue lupine
C:Species: Lupinus angustifolius (narrow-leaved blue lupine)
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 11-Jan-2000
C:Accession: A23617
R:Gillley, G.G.; Inglis, A.S.
FEBS Lett. 195, 235-241, 1986
A:Title: Amino acid sequence of conglutinin delta-2, a sulfur-rich seed protein of Lupinus
A:Reference number: A91458
A:Accession: A23617
A:Molecule type: protein
A:Residues: 1-37 <L11>
C:Superfamily: soybean 2S albumin

Query Match 1.8%; Score 5; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 220 EEEED 224
Db 33 EEEED 37

RESULT 39
C82482
hypothetical protein VCA025; [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: C82482

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, J.;
charlson, D.; Ermlaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Draqui, I.; Seller,
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: C82482
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-38 <HE1>
A:Cross-references: GB:AF004365; GB:AF003853; NID:g9657642; PIDN:AAF96162.1; GSPDB:C
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0251
A:Map position: 2

Query Match 1.8%; Score 5; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 124 LLSVI 128
Db 4 LLSVI 8

RESULT 40
A44862
microtubule associated protein (MAP) homolog - Trypanosoma cruzi (fragment)
C:Species: Trypanosoma cruzi
C>Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1998
C:Accession: A44862
R:Kerner, N.; Liegeard, P.; Levin, M.J.; Hontebeyrie-Joskowicz, M.
Exp. Parasitol. 73, 451-459, 1991
A:Title: Trypanosoma cruzi: antibodies to a MAP-like protein in chronic Chagas' dis
A:Reference number: A44862; MUID:92070444; PMID:1959572
A:Accession: A44862
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-38 <KER>
A:Cross-references: GB:S68286; NID:g239898; PID:g239899
A:Experimental source: epimastigotes, Tulahuen 2 strain
A>Note: sequence extracted from NCBI backbone (NCBI:68286, NCBI:68287)

Query Match 1.8%; Score 5; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 220 EEEED 224
Db 4 EEEED 8

Search completed: April 10, 2003, 10:42:20
Job time : 24.9012 secs

```


GenCore version 5.1.4_p5_4573
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on: April 10, 2002, 10:33:10 ; Search time 29.5271 Seconds
        (without alignments)
        1272.618 Million cell updates/sec
```

```

Title: US-09-330-026-2
Perfect score: 282
Sequence: 1 MAQQMTSSKAI MFLKSL.....AMAPQKP: HSGWNT HSSC 282

```

Scoring table: `OLIGO`
`Gapop 60.0` , `Gapext 60.0`

Searched: 90847 seqs, 733250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 507025

Minimum	DB	seq	length:	0
Maximum	DB	seq	length:	70

Post-processing: Listing first 135 summaries

```

Database :
A_Geneseq_1c1002.*
1: /SID52/qcdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SID52/qcdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SID52/qcdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SID52/qcdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SID52/qcdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SID52/qcdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SID52/qcdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SID52/qcdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SID52/qcdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SID52/qcdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SID52/qcdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SID52/qcdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SID52/qcdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SID52/qcdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SID52/qcdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SID52/qcdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SID52/qcdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SID52/qcdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SID52/qcdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SID52/qcdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SID52/qcdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SID52/qcdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SID52/qcdata/geneseq/geneseqp-emb1/AA2002.DAT.*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	9	3	2	48	21	Zea mays protein f
2	8	2	8	65	22	Peptide #11674 enc
3	8	2	8	65	22	Protein #9048 enc
4	8	2	8	65	22	Human brain expres
5	8	2	8	65	22	Human bone marrow
6	8	2	8	65	22	Peptide #8231 enc
7	8	2	8	65	22	Peptide #12159 enc
8	8	2	8	65	23	Human peptide enc
9	7	2	5	17	18	N-terminal sequenc
10	7	2	5	22	14	Lipopolysaccharide

84	6	2.1	47	22	AAW5685C	Human brain expres	PN	EPI033405-A2.
85	6	2.1	47	22	AAW7064	Peptide #3498 enco	XX	
86	6	2.1	47	22	AAW29558	Peptide #3595 enco	PD	06-SEP-2000.
87	6	2.1	47	22	AAW04766	Peptide #3448 enco	XX	
88	6	2.1	48	17	AAW06970	Modified hsp60 gen	PF	25-FEB-2000; 2000EP-0301439.
89	6	2.1	49	22	ABH39516	Peptide #7022 enco	XX	
90	6	2.1	49	22	ABH24261	Protein #6260 enco	XX	
91	6	2.1	49	22	AAW60204	Human brain expres	PR	25-FEB-1999; 99US-0121825.
92	6	2.1	49	22	AAW72830	Human bone marrow	PR	05-MAR-1999; 99US-0123180.
93	6	2.1	49	22	AAW33060	Peptide #7097 enco	PR	09-MAR-1999; 99US-0123548.
94	6	2.1	49	23	ABG42659	Human peptide enco	PR	23-MAR-1999; 99US-0125788.
95	6	2.1	50	21	AAW56577	Human peptide thalia	PR	25-MAR-1999; 99US-0126264.
96	6	2.1	50	22	AAW89370	Arabidopsis thalia	PR	29-MAR-1999; 99US-0126785.
97	6	2.1	51	21	AAW26832	Zea mays protein f	PR	01-APR-1999; 99US-0127462.
98	6	2.1	51	23	ABF09655	Human OREX protein	PR	06-APR-1999; 99US-0128234.
99	6	2.1	52	22	ABH27718	Human peptide #369	PR	08-APR-1999; 99US-0128714.
100	6	2.1	52	22	ABH32889	Peptide #395 enco	PR	16-APR-1999; 99US-0129845.
101	6	2.1	52	22	ABH18369	Protein #368 enco	PR	19-APR-1999; 99US-0130077.
102	6	2.1	52	22	AAW53691	Human brain expres	PR	21-APR-1999; 99US-0130449.
103	6	2.1	52	22	AAW66074	Human bone marrow	PR	23-APR-1999; 99US-0130510.
104	6	2.1	52	22	AAW13944	Peptide #378 enco	PR	23-APR-1999; 99US-0130891.
105	6	2.1	52	22	AAW26350	Peptide #387 enco	PR	28-APR-1999; 99US-0131449.
106	6	2.1	52	22	AAW01687	Peptide #369 enco	PR	30-APR-1999; 99US-0132048.
107	6	2.1	52	23	ABG35723	Human peptide enco	PR	30-APR-1999; 99US-0132407.
108	6	2.1	53	21	AAW56576	Arabidopsis thalia	PR	04-MAY-1999; 99US-0132484.
109	6	2.1	53	22	ABG03003	Novel human diagno	PR	05-MAY-1999; 99US-0132485.
110	6	2.1	53	23	ABF06959	Human OREX protein	PR	06-MAY-1999; 99US-0132486.
111	6	2.1	54	21	AAW01928	Human secreted pro	PR	07-MAY-1999; 99US-0132487.
112	6	2.1	54	22	ABH38929	Peptide #6435 enco	PR	07-MAY-1999; 99US-0132863.
113	6	2.1	54	22	ABH40112	Peptide #7618 enco	PR	11-MAY-1999; 99US-0134256.
114	6	2.1	54	22	ABH23917	Protein #5916 enco	PR	14-MAY-1999; 99US-0134218.
115	6	2.1	54	22	AAW59580	Human brain expres	PR	14-MAY-1999; 99US-0134219.
116	6	2.1	54	22	AAW72150	Human bone marrow	PR	14-MAY-1999; 99US-0134221.
117	6	2.1	54	22	AAW73555	Peptide #5896 enco	PR	14-MAY-1999; 99US-0134370.
118	6	2.1	54	22	AAW9462	Peptide #6449 enco	PR	18-MAY-1999; 99US-0134376.
119	6	2.1	54	22	AAW32412	Peptide #7785 enco	PR	19-MAY-1999; 99US-0134941.
120	6	2.1	54	22	AAW33748	Human peptide enco	PR	20-MAY-1999; 99US-0135124.
121	6	2.1	54	23	ABG41964	Human peptide enco	PR	21-MAY-1999; 99US-0135153.
122	6	2.1	54	23	ABG43418	Human mucosal lymph	PR	24-MAY-1999; 99US-0135629.
123	6	2.1	55	16	AAW82635	Propionibacterium	PR	25-MAY-1999; 99US-0136021.
124	6	2.1	55	22	AAW4030	Human brain expres	PR	27-MAY-1999; 99US-0136392.
125	6	2.1	55	22	AAW60573	Human bone marrow	PR	28-MAY-1999; 99US-0136782.
126	6	2.1	55	22	AAW73237	Peptide #8451 enco	PR	01-JUN-1999; 99US-0137222.
127	6	2.1	55	22	AAW33084	Novel human diagno	PR	03-JUN-1999; 99US-0137528.
128	6	2.1	55	23	ABG43084	Peptide #8541 enco	PR	04-JUN-1999; 99US-0137502.
129	6	2.1	56	22	ABG26499	Human brain expres	PR	07-JUN-1999; 99US-0137724.
130	6	2.1	57	22	ABH41035	Peptide #8541 enco	PR	08-JUN-1999; 99US-0138094.
131	6	2.1	57	22	AAW61890	Human bone marrow	PR	10-JUN-1999; 99US-0138540.
132	6	2.1	57	22	AAW74692	Peptide #845 enco	PR	10-JUN-1999; 99US-0138847.
133	6	2.1	57	22	AAW34808	Oropouche NP prote	PR	14-JUN-1999; 99US-0139119.
134	6	2.1	58	20	AAW99092	Novel human diagno	PR	16-JUN-1999; 99US-0139452.
135	6	2.1	58	22	ABG19930		PR	17-JUN-1999; 99US-0139492.
							PR	18-JUN-1999; 99US-0139454.
							PR	18-JUN-1999; 99US-0139455.
							PR	18-JUN-1999; 99US-0139456.
							PR	18-JUN-1999; 99US-0139457.
							PR	18-JUN-1999; 99US-0139458.
							PR	18-JUN-1999; 99US-0139459.
							PR	18-JUN-1999; 99US-0139460.
							PR	18-JUN-1999; 99US-0139461.
							PR	18-JUN-1999; 99US-0139462.
							PR	18-JUN-1999; 99US-0139463.
							PR	18-JUN-1999; 99US-0139750.
							PR	18-JUN-1999; 99US-0139763.
							PR	21-JUN-1999; 99US-0139817.
							PR	22-JUN-1999; 99US-0139859.
							PR	23-JUN-1999; 99US-0140353.
							PR	23-JUN-1999; 99US-0140354.
							PR	24-JUN-1999; 99US-0140695.
							PR	28-JUN-1999; 99US-0140823.
							PR	29-JUN-1999; 99US-0140991.
							PR	30-JUN-1999; 99US-0141287.
							PR	01-JUL-1999; 99US-0141842.

ALIGNMENTS

RESULT :
 AAG19155
 ID AAG19155 standard; Protein: 48 AA.

AC AAG19155;

DT 17-OCT 2000 (first entry)

DE Zea mays protein fragment SEQ ID NO: 20850.

XX Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; core.

OS Zea mays subsp. mays.

XX

PR 01-JUL-1999; 990S-0142154.
 PR 02-JUL-1999; 990S-0142055.
 PR 06-JUL-1999; 990S-0142330.
 PR 08-JUL-1999; 990S-0142833.
 PR 09-JUL-1999; 990S-0142920.
 PR 12-JUL-1999; 990S-0142977.
 PR 13-JUL-1999; 990S-0143542.
 PR 14-JUL-1999; 990S-0143624.
 PR 15-JUL-1999; 990S-0144005.
 PR 16-JUL-1999; 990S-0144085.
 PR 19-JUL-1999; 990S-0144325.
 PR 19-JUL-1999; 990S-0144331.
 PR 19-JUL-1999; 990S-0144332.
 PR 19-JUL-1999; 990S-0144333.
 PR 19-JUL-1999; 990S-0144334.
 PR 19-JUL-1999; 990S-0144335.
 PR 20-JUL-1999; 990S-0144332.
 PR 20-JUL-1999; 990S-0144632.
 PR 20-JUL-1999; 990S-0144894.
 PR 21-JUL-1999; 990S-0144814.
 PR 21-JUL-1999; 990S-0145086.
 PR 21-JUL-1999; 990S-0145088.
 PR 22-JUL-1999; 990S-0145085.
 PR 22-JUL-1999; 990S-0145087.
 PR 22-JUL-1999; 990S-0145089.
 PR 22-JUL-1999; 990S-0145192.
 PR 23-JUL-1999; 990S-0145145.
 PR 23-JUL-1999; 990S-0145218.
 PR 23-JUL-1999; 990S-0145224.
 PR 26-JUL-1999; 990S-0145215.
 PR 27-JUL-1999; 990S-0145913.
 PR 27-JUL-1999; 990S-0145918.
 PR 27-JUL-1999; 990S-0145919.
 PR 28-JUL-1999; 990S-0145951.
 PR 02-AUG-1999; 990S-0146386.
 PR 02-AUG-1999; 990S-0146388.
 PR 03-AUG-1999; 990S-0146389.
 PR 03-AUG-1999; 990S-0147038.
 PR 04-AUG-1999; 990S-0147204.
 PR 04-AUG-1999; 990S-0147302.
 PR 05-AUG-1999; 990S-0147192.
 PR 05-AUG-1999; 990S-0147260.
 PR 06-AUG-1999; 990S-0147303.
 PR 06-AUG-1999; 990S-0147416.
 PR 09-AUG-1999; 990S-0147453.
 PR 09-AUG-1999; 990S-0147935.
 PR 10-AUG-1999; 990S-0148171.
 PR 11-AUG-1999; 990S-0148319.
 PR 12-AUG-1999; 990S-0148341.
 PR 13-AUG-1999; 990S-0148565.
 PR 16-AUG-1999; 990S-0149368.
 PR 17-AUG-1999; 990S-0149275.
 PR 18-AUG-1999; 990S-0149426.
 PR 20-AUG-1999; 990S-0149722.
 PR 20-AUG-1999; 990S-0149723.
 PR 20-AUG-1999; 990S-0149929.
 PR 23-AUG-1999; 990S-0149902.
 PR 23-AUG-1999; 990S-0149930.
 PR 25-AUG-1999; 990S-0150566.
 PR 26-AUG-1999; 990S-0150884.
 PR 27-AUG-1999; 990S-0151065.
 PR 27-AUG-1999; 990S-0151066.
 PR 27-AUG-1999; 990S-0151086.
 PR 30-AUG-1999; 990S-0151303.
 PR 31-AUG-1999; 990S-0151438.
 PR 01-SEP-1999; 990S-0151930.
 PR 07-SEP-1999; 990S-0152363.
 PR 10-SEP-1999; 990S-0153070.
 PR 13-SEP-1999; 990S-0153758.
 PR 15-SEP-1999; 990S-0154018.
 PR 16-SEP-1999; 990S-0154039.

PR 20-SEP-1999; 990S-0154779.
 PR 22-SEP-1999; 990S-0155139.
 PR 23-SEP-1999; 990S-0155486.
 PR 24-SEP-1999; 990S-0155659.
 PR 28-SEP-1999; 990S-0156458.
 PR 29-SEP-1999; 990S-0156596.
 PR 04-OCT-1999; 990S-0157117.
 PR 05-OCT-1999; 990S-0157753.
 PR 06-OCT-1999; 990S-0157865.
 PR 07-OCT-1999; 990S-0158029.
 PR 08-OCT-1999; 990S-0158232.
 PR 12-OCT-1999; 990S-0158369.
 PR 13-OCT-1999; 990S-0159293.
 PR 13-OCT-1999; 990S-0159294.
 PR 13-OCT-1999; 990S-0159295.
 PR 14-OCT-1999; 990S-0159329.
 PR 14-OCT-1999; 990S-0159330.
 PR 14-OCT-1999; 990S-0159331.
 PR 14-OCT-1999; 990S-0159637.
 PR 14-OCT-1999; 990S-0159638.
 PR 18-OCT-1999; 990S-0159584.
 PR 21-OCT-1999; 990S-0160741.
 PR 21-OCT-1999; 990S-0160767.
 PR 21-OCT-1999; 990S-0160768.
 PR 21-OCT-1999; 990S-0160770.
 PR 21-OCT-1999; 990S-0160814.
 PR 21-OCT-1999; 990S-0160815.
 PR 22-OCT-1999; 990S-0160980.
 PR 22-OCT-1999; 990S-0160981.
 PR 22-OCT-1999; 990S-0160989.
 PR 25-OCT-1999; 990S-0161404.
 PR 25-OCT-1999; 990S-0161405.
 PR 26-OCT-1999; 990S-0161406.
 PR 26-OCT-1999; 990S-0161359.
 PR 26-OCT-1999; 990S-0161360.
 PR 26-OCT-1999; 990S-0161361.
 PR 28-OCT-1999; 990S-0161920.
 PR 28-OCT-1999; 990S-0161992.
 PR 28-OCT-1999; 990S-0161993.
 PR 29-OCT-1999; 990S-0162142.

Query Match 3.2%; Score 9; DB 21; Length 48;

Best Local Similarity 100.0%; Pred. No. 0.25;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 VCIS:LHPP 100

DB 3: VCIS:LHPP 39

RESULT 2

ABR44168

ID ABR44168 standard; Peptide; 65 AA.

XX ABR44168;

AC ABR44168;

XX 04-FEB-2002 (first entry)

DI Peptide #11674 encoded by human foetal liver single exon probe.

DE Human; foetal liver; gene expression: single exon nucleic acid probe.

XX Homo sapiens.

OS WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

```

PR 04-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
PT
XX
XX Claim 27: SEQ ID NO 34803: 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 AA:
      Query Match      2.8%; Score 8; DB 22; Length 65;
      Best Local Similarity 100.0%; Pred. No. 3.5;
      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 DEEEDAD 226
DB 45 DEEEDAD 52
      |||||
      |||||

RESULT 3
ABH27049
ID ABH27049 standard; Protein: 65 AA.
XX
AC ABH27049;
XX
XX
DT 23-JAN-2002 (first entry)
XX
XX Protein #9048 encoded by probe for measuring heart cell gene expression.
XX
XX Human; gene expression; heart; microarray; vascular system;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG 2002.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX

```

```

PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
XX Claim 15; SEQ ID NO 28819: 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX AB21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease.
XX Hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 AA:
      Query Match      2.8%; Score 8; DB 22; Length 65;
      Best Local Similarity 100.0%; Pred. No. 3.5;
      Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 DEEEDAD 226
DB 45 DEEEDAD 52
      |||||
      |||||

RESULT 4
AAM65195
ID AAM65195 standard; Protein: 65 AA.
XX
XX AAM65195;
XX
XX
DT 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 37300.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 37300; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX

```

CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX SQ Sequence 65 AA;

Query Match 2.8%; Score 8; DB 22; Length 65;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 DEEEDAD 226
 |||||
 Db 45 DEEEDAD 52

RESULT 5

AAW77903
 ID AAW77903 standard; Protein: 65 AA.

XX AC AAW77903;

XX DT 26-NOV-2001 (first entry)

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 38209.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;

XX KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PE 30-JAN-2001; 2001WO-US06668.

XX PR 04-FEB-2000; 2000US-0183312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PL Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human bone marrow

XX PS Example 4; SEQ ID NO: 38209; 658bp * Sequence Listing: English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC bone marrow. They can be used to measure gene expression in bone marrow

XX CC samples, which may enable the improved diagnosis and treatment of cancers

XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a

XX CC protein encoded by one of the probes of the invention.

XX SQ Sequence 65 AA;

RESULT 6

AAW21797
 ID AAW21797 standard; Protein: 65 AA.

XX AC AAW21797;

XX DT 12-OCT-2001 (first entry)

XX DE Peptide #8231 encoded by probe for measuring cervical gene expression.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US060670.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PL Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-488901/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human cervical epithelial cells

XX PS Claim 27; SEQ ID NO 26623; 487bp; English.

XX CC The present invention relates to human single exon nucleic acid probes

XX CC (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded

XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX CC can be used to produce a single exon microarray, which can be used for

XX CC measuring human gene expression in a sample derived from human cervical

XX CC epithelial cells. By measuring gene expression, the probes are therefore

XX CC useful in grading and/or staging of diseases of the cervix, notably

XX CC cervical cancer.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 65 AA;

Query Match 2.8%; Score 8; DB 22; Length 65;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 DEEEDAD 226
 |||||
 Db 45 DEEEDAD 52

RESULT 7

AAW38122
 ID AAW38122 standard; Protein: 65 AA.

XX AC AAW38122;

XX DT 17-OCT-2001 (first entry)

XX DE Peptide #12159 encoded by probe for measuring placental gene expression.

XX Probe: microarray; human; Placenta: antenatal diagnosis;
 KW genetic disorder;
 XX Homo sapiens.
 XX W0200157272-A2.
 XX 09-AUG-2003.
 XX 30-JAN-2003; 2001WO-US006663.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX Claim 27; SEQ ID No 38391; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP;
 CC see AA131315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX Sequence 65 AA;
 SQ
 Query Match 2.8%; Score 8; DB 22; Length 65;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 219 DEEFEDAD 226
 III-III
 DB 45 DEEFEDAD 52
 RESULT 8
 ABG46923
 ID ABG46923 standard; Peptide; 65 AA.
 AC ABG46923;
 XX
 XX 10-AUG-2002 (first entry)
 DT
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 36588.
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary alveolar proteinosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW primary ciliary dyskinesia; fibrocystic pulmonary dysplasia;
 KW hyaline membrane disease.
 XX Homo sapiens.
 XX W0200186003-A2.
 XX

PD 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US006665.
 XX 04-FEB-2000; 2000US-180312P.
 XX 26-MAY-2000; 2000US-207456P.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-234687P.
 XX 27-SEP-2000; 2000US-236359P.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX Claim 27; SEQ ID No 36588; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 65 AA;
 SQ
 Query Match 2.8%; Score 8; DB 23; Length 65;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 219 DEEFEDAD 226
 III-III
 DB 45 DEEFEDAD 52

```

RESULT 9
AAW23427
ID AAW23427 standard; Peptide: 17 AA.
XX AC AAW23427;
XX
XX DI 08-APR-1998 (first entry)
XX
XX DE N-terminal sequence of large chain of albumin 3 protein.
XX
XX KW Soybean albumin 3; seed storage protein; SSP; transgenic seed;
XX transgenic soybean plant; animal feed production; N-terminus.
XX
XX OS Glycine max.
XX
XX FH Key Location/Qualifiers
FH Misc-difference 16 /note- "unspecified amino acid"
FH Misc-difference 17 /note- "unspecified amino acid"
XX
XX FN W09735023-AZ.
XX
XX PD 25-SEP-1997.
XX
XX PF 19-MAR-1997; 9/WO-US04405.
XX
XX PR 20-MAR-1996; 9GUS-06:8911.
XX
XX PA (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Coughlan S, Hastings C, Hu D, Jung R;
XX
XX DK WPI: 1997-480228/44.
XX
XX PI Increasing the nutritional quality of soybean - by genetically
PI increasing the amount of lysine, methionine and/or cysteine in
PI albumin proteins
XX
XX PS Example 1: Fig 1: 62pp; English.
XX
XX CC This sequence represents the N-terminal fragment of soybean albumin 3
XX protein. DNA encoding the full length protein is a DNA molecule of the
XX invention. The DNA molecules of the invention comprise a preselected DNA
XX segment encoding a seed storage protein (SSP). (-) is used to produce
XX transgenic seeds and plants, especially soybean plants that have
XX increased levels of lysine and also increased levels of methionine and/or
XX cysteine in albumin protein type 1 and/or 3. Increasing the levels of
XX these amino acids increases the nutritional value of soybean produced.
XX This is especially useful for producing animal feeds. The amount of
XX lysine in the seed is increased by 5-10 %, the amount of methionine and
XX cysteine is increased by 10-30%
XX
XX SQ Sequence 17 AA;
XX
Query Match 2.5%; Score 7; DB 18; Length 17;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 DEDEEE 223
DB 4 DEDEEE 10

RESULT 13
AAK51952
ID AAK51952 standard; Protein: 22 AA.
XX AC AAK51952;
XX
XX DT 24-MAY-1994 (first entry)
XX
Lipopolysaccharide induced protein (peptide fragment).
XX
XX DE Macrophage; induced; lipo-polysaccharide; antitumour;
XX antiinflammatory; trypanocidal agent; antibody; cell proliferation;
XX activation; cytotoxicity.
XX
XX OS Homo sapiens.
XX
XX PN W09322437-A.
XX
XX PD 11-NOV-1993.
XX
XX PE 28-APR-1993; 93WO-EP01022.
XX
XX PR 30-APR-1992; 92EP-0401231.
XX
XX PA (INNO-) INNOGENETICS NV SA.
XX
XX PI Devos K, Franssen L, Van De Voorde A, Van Heuverswyn B;
XX
XX DR WPI: 1993-368796/46.
XX
XX PI New polypeptide induced in macrophage(s) by lipo-polysaccharide -
PI useful e.g. as antitumour, antiinflammatory or trypanocidal
PI agent, also related nucleic acid, antibodies, anti-sense cpds.
PI etc.
XX
XX PS Claim 3; Page 66; 108pp; English.
XX
XX CC The polypeptide induced in macrophages by lipopolysaccharide
XX stimulates cell proliferation (esp. when costimulated with IL-4)
XX promote activation, cytotoxicity, and mobilisation of TAC cells;
XX promote recruitment of suppressive peritoneal exudate cells;
XX promote generation of immunocompetent lymph node cells (TNC) and
XX have trypanocidal and trypanolytic activity. The human and murine
XX sequences are given in (AA051543-45). Peptide fragments able to
XX generate antibodies are given in (AAR51951-61)
XX
XX SQ Sequence 22 AA;
XX
Query Match 2.5%; Score 7; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 PNTFSPA 140
DB 16 PNTFSPA 22

RESULT 11
ABR05466
ID ABR05466 standard; peptide: 31 AA.
XX AC ABR05466;
XX
XX DT 17-APR-2002 (first entry)
XX
XX DE Peptide with growth hormone production increasing activity SEQ ID:21.
XX
XX KW Medical; hypertension; osteoporosis; dwarfism; hypotensive; pain-killing;
XX growth hormone production inhibiting; fat accumulation inhibiting;
XX blood calcium increasing; gastric juice secretion inhibiting;
XX prostaglandin E2 production inhibiting; osteoblast growth promoting;
XX growth hormone production promoting.
XX
XX OS Synthetic.
XX
XX PN JP2001335596-A.
XX
XX PD 04-DEC-2001.
XX
XX PF 24-MAY-2000; 2000JP-0152459.

```


PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226581.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227029.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231988.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233053.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234224.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241858.
 PR 20-OCT-2000; 2000US-0241859.
 PR 20-OCT-2000; 2000US-0241859.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0245474.
 PR 08-NOV-2000; 2000US-0245475.
 PR 08-NOV-2000; 2000US-0245476.
 PR 08-NOV-2000; 2000US-0245477.
 PR 08-NOV-2000; 2000US-0245478.
 PR 08-NOV-2000; 2000US-0245523.
 PR 08-NOV-2000; 2000US-0245524.
 PR 08-NOV-2000; 2000US-0245526.
 PR 08-NOV-2000; 2000US-0245526.
 PR 08-NOV-2000; 2000US-0245527.
 PR 08-NOV-2000; 2000US-0245528.
 PR 08-NOV-2000; 2000US-0245532.

PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Harash SC, Ruben SM;
 PI
 XX WPI; 2001-483426/52.
 DR N-PSDB; AAK59054.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Claim 11; SEQ ID NO 13876; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK64702 to AAK64702. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK67694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 XX Sequence 47 AA;

Query Match 2.5% Score 7; DB 22; Length 47;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 VEGFR1: 31
 ID AAW79542 standard; Protein: 51 AA.
 DB 19 VEGFR1: 25

RESULT 14

AAW79542
 ID AAW79542 standard; Protein: 51 AA.
 AC AAW79542;

ET 11-JAN-1999 (first entry)

DE Adenovirus serotype 2 hypervariable region HVRI.

KW Adenovirus serotype 2; Ad2; hexon; coat protein; gene therapy;
 vector.

OS Mastadenovirus h2.

XX W09840509-A1.

PN 17-SEP-1998.

PF 13-MAR-1998: 98W0-US05033.

PR 13-MAR-1997: 97US-0816345.

PA (CORR) CORNELL RES FOUND INC.
 (GENV) GENVEC INC.

XX Crystal R3, Falck Pedersen E, Gal J, Kovacs I;
 Wickham LC;

DR WPI: 1998-536738/43.
 N PSDB: AAV61504.

PI Chimeric adenovirus coat protein - useful in, e.g. vector for gene
 transfer to treat inherited genetic diseases

PS Claim 7: Page 76-77: 112pp; English.

CC This is the amino acid sequence of hypervariable region HVRI of the
 hexon protein (see AAW79538) of adenovirus serotype 2 (Ad2). It is
 encoded by a claimed DNA sequence (see AAV61504). The invention
 provides a chimeric adenoviral coat protein, particularly a chimeric
 adenovirus hexon protein, that has a decreased ability of inability
 to be recognised by a neutralising antibody directed against the
 corresponding wild-type adenovirus coat protein. The chimeric
 adenoviral coat protein has a non-native amino acid sequence,
 especially comprising a deletion of an internal hexon protein
 sequence, preferably a hypervariable region or entire loop. DNA
 sequences (see AAV61502-23) encoding claimed Ad2 and Ad5 chimeric
 coat proteins (see AAW79540-61) are claimed. Also claimed are an
 adenovirus vector that comprises the chimeric adenovirus coat
 protein, a method of genetically modifying a cell by contacting it
 with the vector, and a host cell that comprises the chimeric
 adenovirus coat protein. The vector can be used for gene transfer,
 for the treatment of inherited diseases. It can also be used to
 render certain cells susceptible to the killing action of certain
 drugs, or to study the effects of expression of specific genes in a
 given cell or tissue in vitro or in vivo.

XX Sequence 51 AA;

Query Match 2.5%; Score 7; DB 19; Length 51;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDDEED 224

DB 10 EDDEED 16

RESULT 15

AAO04342
 ID AAO04342 standard; Protein: 60 AA.

XX AAO04342;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 18234.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 tissue growth factor; immunomodulatory; cancer; leukaemia;
 nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

XX W0200164835-A2.

PN 07-SEP-2001.

PF 26-FEB-2001: 2001WO-US04927.

PR 28-FEB-2000: 2000US-0515126.

PR 18-MAY-2000: 2000US-0577409.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI: 2001-514838/56.
 N PSDB: AAI84273.

PT Isolated nucleic acids and polypeptides, useful for preventing
 diagnosing and treating e.g. leukaemia, inflammation and immune
 disorders

PS Claim 20: SEQ ID NO 18234; 1399pp - Sequence listing; English.

CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 cytokine, cell proliferation or cell differentiation or which may induce
 production of other cytokines in other cell populations. The
 polynucleotides and polypeptides are useful in gene therapy, vaccines or
 peptide therapy. The polypeptides have various cytokine-like activities,
 e.g. stem cell growth factor activity, haematopoiesis regulating
 activity, tissue growth factor activity, immunomodulatory activity and
 activin/inhibin activity and may be useful in the diagnosis and/or
 treatment of cancer, leukaemia, nervous system disorders, arthritis and
 inflammation.

CC Note: The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIPO
 at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 60 AA;

Query Match 2.5%; Score 7; DB 22; Length 60;

Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 DDEDEFF 222

DB 43 DDEDEFF 49

RESULT 16

AAE18293
 ID AAE18293 standard; peptide: 7 AA.

XX AAE18293;

DT 07-MAY-2002 (first entry)
 DE Antigenic peptide #2 used for purification of a target protein.
 XX
 KW Bovine; recombinant protein; larvae expression system; membrane protein;
 KW transport protein; cardiac sodium-calcium exchange protein; Na-K ATPase;
 KW NCX1; cystic fibrosis transmembrane conductance regulator; CFTR; vaccine;
 KW channel forming protein; junctional protein; connexin 32;
 KW antigenic peptide.
 XX
 OS Unidentified.
 XX WO200206464-A2.
 PN 24-JAN-2002.
 PD 09-JUL-2002; 2001WO-US21606.
 XX 13-JUL-2000; 2000US-218125P.
 PR (OMOR) UNIV M'ISSOURI.
 XX Title CC, Price EW;
 XX WPI: 2002-171806/22.
 DR Producing recombinant proteins e.g. membrane, transport and channel
 XX forming proteins in larvae expression system, by infecting larvae with
 PT vector having a sequence encoding recombinant fusion protein with
 PT affinity tag -
 XX Disclosure: Page 9; 40pp; English.
 XX The patent discloses methods of producing recombinant proteins in larvae
 CC expression system, by infecting the larvae with vector having a sequence
 CC encoding recombinant fusion protein with affinity tag. The methods are
 CC useful for producing recombinant protein, preferably membrane proteins,
 CC transport proteins such as NCX1 (cardiac sodium-calcium exchange protein)
 CC or Na-K ATPase, channel forming proteins such as cystic fibrosis trans-
 CC membrane conductance regulator (CFTR), junctional protein (connexin 32),
 CC receptor, cytoskeletal and other membrane associated proteins. They are
 CC also useful for producing prostate specific membrane antigens and sodium
 CC phosphate co-transporters from kidney. The methods are also useful for
 CC producing recombinant fusion proteins in large quantities that are both
 CC highly homogeneous and biologically active. The recombinant proteins
 CC produced by the methods of the invention can be included as part of a
 CC pharmaceutical, nutritional, drug or vaccine composition. The present
 CC sequence is an antigenic peptide which is used for the purification of
 CC a target protein. The immunogenic tag allows the protein to which it is
 CC attached to be purified based upon its affinity for an antibody.
 XX
 SQ Sequence 7 AA:
 Query Match 2.1%; Score 6; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 DLYDDD 214
 DB 1 DLYDDD 6
 RESULT 17
 ID AAY77718
 XX AAY77718 standard; peptide: 8 AA.
 AC AAY77718:
 XX
 DT 12-MAY-2000 (first entry)
 DE Xpress peptide epitope.
 XX Cell surface receptor; luminescence; protein internalization;
 KW

drug discovery; screening assay; epitope.
 KW Synthetic.
 OS WO200003246-A2.
 PN 20-JAN-2000.
 PD 13-JUL-1999; 99WO-US15870.
 XX 13-JUL-1998; 98US-0092671.
 PR (CEL.-) CELLOMICS INC.
 XX Rubin RA, Giuliano KA, Gough A, Dunlay T;
 XX WPI: 2000-171170/15.
 DR Automated screening method for identifying compounds which induce cell
 XX surface receptor internalization, useful for drug discovery -
 XX Example 6; Page 67; 148pp; English.
 XX The invention relates to a method for identifying compounds which
 CC inhibit internalization of cell surface receptors. Provided are an array
 CC of locations, each containing cells with a cell surface receptor
 CC protein, that are treated with a test compound. The protein is
 CC luminescently labeled or contacted with a luminescently labeled cell
 CC before or after test compound treatment. Any luminescence produced is
 CC converted into digital data and automatically analysed to determine if
 CC the test compound induced the protein internalization. The novel method
 CC is used to screen for compounds which modulate cell surface receptor
 CC protein internalization, this can be used in drug discovery, to test
 CC compound efficacy in living biological systems. The assay method is
 CC automated and compact. It has high throughput and uses smaller volumes of
 CC reagents and test compounds. Sequences AAY77704-718 represent examples of
 CC peptide epitope tags used in the course of the invention.
 XX
 SQ Sequence 8 AA:
 Query Match 2.1%; Score 6; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 DLYDDD 214
 DB 1 DLYDDD 6
 RESULT 18
 ID AAB97365
 XX AAB97365 standard; peptide: 8 AA.
 AC AAB97365:
 XX
 DT 15-AUG-2001 (first entry)
 DE Xpress epitope used in dual labelled receptor construction.
 XX Automated measurement; cell viability; epitope tag; luminescence;
 KW G-protein coupled receptor; high content screen.
 XX Unidentified.
 XX WO200135072-A2.
 PN 17-MAY-2001.
 PD C9-NOV-2000; 2000WO-US30896.
 PF C9-NOV-1999; 99CS-0164353.
 PR 18-JAN-2000; 2000US-0176504.
 XX

PA (CELL-) CELLOMICS INC.
 XX Ghosh RN, Debiasio R, Chen Y, Bellutta P, Giuliano K, Pasley JW;
 XX WPI: 2001-329169/34.
 XX Automated measurement of cell viability, involves contacting cells with
 PT luminescent reporter molecule, imaging cells to get signals, converting
 PT signals into digital data and using data to measure viable cell -
 XX Example 6; Page 52; 155pp; English.
 XX This invention relates to a method for the automated measurement of cell
 CC viability. The method involves contacting cells with luminescent
 CC reporter molecules, imaging cells to get signals, and converting the
 CC signals into digital data which can be used as a measurement of cell
 CC viability. Included in the invention is a computer readable storage
 CC medium comprising a programme which causes the method of the invention
 CC to be activated. The method is useful for cell state identification in
 CC cells. The method is also useful for drug discovery. An example of the
 CC invention relates to the use of inserted sequences and their ligands for
 CC high content screens incorporating dual labelled receptors. The present
 CC sequence represents an epitope tag used to label one end of a G-protein
 CC coupled receptor (GPCR). The intracellular and extracellular domains of
 CC the GPCR are distinctly labelled so that using the method of the
 CC invention the extent of internalisation of the receptor can be measured.
 XX Sequence 8 AA;
 SQ
 Query Match 2.1%; Score 6; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 DLYDD 214
 DB 1 DLYDD 6
 RESULT 19
 AAB59869
 ID AAB59869 standard; peptide; 8 AA.
 XX AAB59869;
 AC AAB59869;
 XX 30-MAR-2001 (first entry)
 DT Xpress peptide epitope.
 DE Xpress peptide epitope.
 XX Xpress peptide epitope; macromolecule trafficking; endosomal system;
 KW membrane receptor internalisation.
 XX Unidentified.
 XX W0200075241 A2.
 PN 28-DEC-2000.
 PD 21-JUN-2000; 2000WO-US40250.
 PF 21-JUN-1999; 99US-0140143.
 PR 12-JUL-1999; 99US-0352171.
 PR 11-AUG-1999; 99US-0148360.
 PR 13-DEC-1999; 99US-0170313.
 XX (CELL-) CELLOMICS INC.
 PA Rubin RA, Gough AH, Ghosh RN, Giuliano KA, Dunlay RT;
 XX WPI: 2001-031619/10.
 XX Identifying compounds modulating macromolecule trafficking through
 PT endosomes, using digital data obtained by converting a luminescent
 PI signal from cells contacted with the compound -

XX Example 6; Page 53; 113pp; English.
 XX The present invention relates to an automated method for identifying
 CC compounds that induce or inhibit macromolecule trafficking through an
 CC endosomal system. The method comprises treating cells which possess a
 CC luminescently-tagged macromolecule, with a test compound, and obtaining
 CC luminescent signals from the cells. The signal is converted into digital
 CC data that is used to determine if the test compound has induced or
 CC inhibited the trafficking. The method can also be used to identify the
 CC extent of internalisation of membrane receptors, by fusing a labelled
 CC peptide epitope to the different domains of the receptor e.g. the
 CC extracellular domain and intracellular domain. The present sequence is
 CC one such peptide epitope used in the method of the present invention.
 XX Sequence 8 AA;
 SQ
 Query Match 2.1%; Score 6; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 DLYDD 214
 DB 1 DLYDD 6
 RESULT 20
 AAY11697
 ID AAY11697 standard; peptide; 9 AA.
 XX AAY11697;
 AC AAY11697;
 XX 18-JUN-1999 (first entry)
 DT C. ruminantium major antigenic protein 1 (MAP-1) peptide fragment.
 DE Major antigenic protein 1; MAP-1; GroEL; heat shock protein; treatment;
 XX Cowdria infection; vaccine; medicament; diagnosis.
 KW Synthetic.
 XX Cowdria ruminantium.
 OS W09514233-A1.
 XX 25-MAR-1999.
 PD 18-SEP-1998; 98WO-GB02768.
 PF 18-SEP-1997; 97GB-0019772.
 PR (UYED-) UNIV EDINBURGH.
 XX Sumption KJ;
 PT WPI: 1999-254380/21.
 DR Cowdria ruminantium major antigen protein 1 and GroEL homologue
 PT peptides
 XX Claim 2; Page 38; 48pp; English.
 XX The invention relates to a fragment of the major antigenic protein 1
 CC (MAP-1) or GroEL homologue protein from Cowdria ruminantium which binds
 CC one or more antibodies produced in response to an Cowdria infection. The
 CC MAP-1 peptide and GroEL homologue peptides are used in vaccines or to
 CC raise antibodies for use in medicaments to diagnose, treat or prevent a
 CC Cowdria species infection. The present sequence represents a
 CC C. ruminantium MAP-1 peptide fragment.
 XX Sequence 9 AA;
 SQ
 Query Match 2.1%; Score 6; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 DGVKVP 184
 DB 4 DGVKVP 9

RESULT 21
 AAR82645
 ID AAR82645 standard; Peptide: 10 AA.
 XX
 AC AAR82645;
 XX
 DT 25-FEB-1996 (first entry)
 DE HML-1 alpha-E chain X domain peptide analogue.
 XX
 KW Human mucosal lymphocyte-1 antigen; HML-1; integrin;
 KW cell attachment; T cell; X domain.
 XX
 OS Homo sapiens.
 XX
 PN WC9522610-A1.
 XX
 PD 24-AUG-1995.
 XX
 PF 15-FEB-1995; 95WO-US02044.
 XX
 PR 18-FEB-1994; 94US-0199776.
 XX
 PA (HGHM) BRIGHAM & WOMENS HOSPITAL.
 XX
 PI Brenner MB, Parker CM;
 XX
 DB WPI: 1995-302716/39.
 XX
 PS New isolated integrin alpha sub-unit and peptide(s) - exhibit cell
 PT attachment activity or block activity of intra-epithelial
 PT lymphocytes
 XX
 PS Claim 1; Page 57; 75pp; English.
 XX
 CC Peptides (AAR82638-R82655) based on, or derived from, a highly-charged
 CC fragment (AAR82636) of the X-domain of the human mucosal lymphocyte-1
 CC alpha-E chain (AAR82656) can be attached to a support to provide a
 CC surface that exhibits cell attachment activity for use e.g. in a
 CC prosthetic device or as an affinity matrix for isolating a ligand of the
 CC alpha-E chain, or can be used in assays for cpds. useful e.g. in
 CC modulating the immune system, tumour progression or metastasis,
 CC or in regulating developmental processes.
 XX
 SQ Sequence 10 AA;

Query Match 2.18; Score 6; DB 16; Length 10;
 Best Local Similarity 100.08; Pred. No. 54;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
 DB 5 EDEEE 10

RESULT 22
 AAR82646
 ID AAR82646 standard; Peptide: 10 AA.
 XX
 AC AAR82646;
 XX
 DT 25-FEB-1996 (first entry)
 DE HML-1 alpha-E chain X domain peptide analogue.
 XX
 DT Human mucosal lymphocyte-1 antigen; HML-1; integrin;
 KW

KW cell attachment; T-cell; X domain.
 XX
 OS Homo sapiens.
 XX
 PN WC9522610-A1.
 XX
 PD 24-AUG-1995.
 XX
 PF 15-FEB-1995; 95WO-US02044.
 XX
 PR 18-FEB-1994; 94US-0199776.
 XX
 PA (HGHM) BRIGHAM & WOMENS HOSPITAL.
 XX
 PI Brenner MB, Parker CM;
 XX
 DB WPI: 1995-302716/39.
 XX
 PS New isolated integrin alpha sub-unit and peptide(s) - exhibit cell
 PT attachment activity or block activity of intra-epithelial
 PT lymphocytes
 XX
 PS Claim 1; Page 58; 75pp; English.
 XX
 CC Peptides (AAR82638-R82655) based on, or derived from, a highly-charged
 CC fragment (AAR82636) of the X-domain of the human mucosal lymphocyte-1
 CC alpha-E chain (AAR82656) can be attached to a support to provide a
 CC surface that exhibits cell attachment activity for use e.g. in a
 CC prosthetic device or as an affinity matrix for isolating a ligand of the
 CC alpha-E chain, or can be used in assays for cpds. useful e.g. in
 CC modulating the immune system, tumour progression or metastasis,
 CC or in regulating developmental processes.
 XX
 SQ Sequence 10 AA;

Query Match 2.18; Score 6; DB 16; Length 10;
 Best Local Similarity 100.08; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEE 223
 DB 4 EDEEE 9

RESULT 23
 AAG95494
 ID AAG95494 standard; Peptide: 10 AA.
 XX
 AC AAG95494;
 XX
 DT 18-SEP-2001 (first entry)
 DE Human complementary peptide, SEQ ID NO: 1688.
 XX
 KW Human; complementary peptide; ligand; drug discovery; drug design.
 XX
 OS Homo sapiens.
 XX
 PN W0200142277-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 13-DEC-2000; 2000WO-GB04776.
 XX
 PR 13-DEC-1999; 99GB-0029464.
 XX
 PA (PROCT-) PROTEOM LTD.
 XX
 PI Roberts GW, Heal JR;
 XX
 DB WPI: 2001-408419/43.
 XX
 PT A set of peptide ligands consisting of specific complementary peptides

PT 10 proteins encoded by genes of the human genome, useful in an assay
 PT for screening and identifying of one or more novel peptides which are
 PT drug candidates or pro-drugs -

PS Example 4; Page 284; 646pp; English.

CC The invention relates to a set of complementary peptide ligands
 CC generated from the human genome. The complementary peptides
 CC interact with their relevant target proteins encoded in the human
 CC genome. They can be used as reagents in drug discovery and as lead
 CC ligands to facilitate drug design and development. The present
 CC sequence is a complementary peptide provided in the specification.

XX Sequence 10 AA:

Query Match 2.18; Score 6; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 CYDDDD 232
 DB 2 CYDDDD 7

RESULT 24

AAR82652 15 AAR82652 standard; Peptide; 12 AA.

AC AAR82652;

DT 25-FEB-1996 (first entry)

DE HML-1 alpha-E chain X domain peptide analogue.

XX Human mucosal lymphocyte-1 antigen; HML-1; integrin;

KW cell attachment; T cell; X domain.

XX Synthetic.

XX WO9522610-A1.

XX 24-AUG-1995.

PF 15-FEB-1995; 95WO-US02044.

PR 18-FEB-1994; 94US-0199776.

PA (BCHM) BRIGHAM & WOMENS HOSPITAL.

XX Brenner MB, Parker CM;

XX WPI; 1995-302716/39.

XX New isolated integrin alpha sub-unit and peptide(s) - exhibit cell
 PT attachment activity or block activity of intra epithelial
 PT lymphocytes

PS Claim 1; Page 60; 75pp; English.

XX Peptides (AAR82638-R82655) based on, or derived from, a highly-charged
 CC fragment. (AAR82636) of the X-domain of the human mucosal lymphocyte-1
 CC alpha-E chain (AAR82656) can be attached to a support to provide a
 CC surface that exhibits cell attachment activity for use e.g. in a
 CC prosthetic device or as an affinity matrix for isolating a ligand of the
 CC alpha-E chain, or can be used in assays for cpds. useful e.g. in
 CC modulating the immune system, tumour progression or metastasis,
 CC or in regulating developmental processes.

XX Sequence 12 AA;

Query Match 2.18; Score 6; DB 16; Length 12;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
 DB 6 EDEEEE 11

RESULT 25

AAR82654

ID AAR82654 standard; Peptide; 12 AA.

XX AAR82654;

DT 25-FEB-1996 (first entry)

DE HML-1 alpha-E chain X domain peptide analogue.

XX Human mucosal lymphocyte-1 antigen; HML-1; integrin;

KW cell attachment; T-cell; X domain.

XX Synthetic.

XX WO9522610-A1.

XX 24-AUG-1995.

PF 15-FEB-1995; 95WO-US02044.

PR 18-FEB-1994; 94US-0199776.

PA (BCHM) BRIGHAM & WOMENS HOSPITAL.

XX Brenner MB, Parker CM;

XX WPI; 1995-302716/39.

XX New isolated integrin alpha sub-unit and peptide(s) - exhibit cell
 PT attachment activity or block activity of intra epithelial
 PT lymphocytes

PS Claim 1; Page 61; 75pp; English.

XX Peptides (AAR82638-R82655) based on, or derived from, a highly-charged
 CC fragment. (AAR82636) of the X-domain of the human mucosal lymphocyte-1
 CC alpha-E chain (AAR82656) can be attached to a support to provide a
 CC surface that exhibits cell attachment activity for use e.g. in a
 CC prosthetic device or as an affinity matrix for isolating a ligand of the
 CC alpha-E chain, or can be used in assays for cpds. useful e.g. in
 CC modulating the immune system, tumour progression or metastasis,
 CC or in regulating developmental processes.

XX Sequence 12 AA;

Query Match 2.18; Score 6; DB 16; Length 12;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
 DB 1 EDEEEE 6

RESULT 26

ABB52033

ID ABB52033 standard; Peptide; 13 AA.

XX ABB52033;

DT 08-FEB-2002 (first entry)

DE Human API-8 tryptic digest peptide #1.

XX Human; neuroprotective; nootropic; gene therapy; vaccine;

KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;

KW Alzheimer's Disease-Associated protein isoform: API; tryptic digest;
 XX Expression Reference Protein Isoform: ERPI; proteolysis.
 OS Homo sapiens.
 PN W0200175454-A2.
 PD 11-OCT-2001.
 XX
 PF 03-APR-2001; 2001WO-US0938.
 XX
 PR 03-APR-2000; 2000US-194504P.
 PR 28-NOV-2000; 2000US-253647P.
 XX
 PA (OXFO-) OXFORD GLYCSO SCIENCES UK LTD.
 PA (PIF-) PFIZER INC.
 PI Durham KL, Friedman DL, Herath HMAC, Kimmel LH, Parekh RB;
 PI Potter DM, Rohlfs C, Silber BM, Stiger TR, Sunderland PT;
 PI Townsend RR, White F, Williams SA;
 XX WPI: 2001-639384/73.
 XX
 XX Screening for Alzheimer's disease in a mammal, by making
 PT two dimensional array of a feature whose relative abundance correlates
 PT with disease, and comparing with abundance of the feature in samples of
 PT healthy persons.
 XX
 PS Example: Page 26; 162pp; English.
 XX
 CC The invention relates to methods for the screening, diagnosis and
 CC prognosis of Alzheimer's disease. The methods involve the detection
 CC of Alzheimer's Disease-Associated Features (AFs) and Alzheimer's
 CC Disease-Associated Protein Isoforms (APIs) in cerebrospinal fluid,
 CC serum or plasma. The abundance of the AFs and APIs is then
 CC normalised to an Expression Reference Protein Isoform (ERPI) in
 CC order to determine whether a patient is suffering from, or has
 CC a predisposition to, Alzheimer's Disease. The relative abundance of
 CC the AFs and APIs correlates with the severity of Alzheimer's Disease.
 CC The present sequence is a peptide produced from an API by proteolysis.
 XX
 SQ Sequence 13 AA;
 Query Match 2.1%; Score 6; DB 22; Length 13;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 218 EDEEEE 223
 DB 1 EDEEEE 6
 RESULT 27
 AAR82648
 ID AAR82648 standard; Peptide; 14 AA.
 XX
 AC AAR82648;
 XX
 DI 25-FEB-1996 (first entry)
 XX
 DE HML-1 alpha-E chain X domain peptide analogue.
 XX
 KW Human mucosal lymphocyte-1 antigen; HML-1; integrin;
 KW cell attachment; T-cell; X domain.
 XX
 OS Homo sapiens.
 XX
 PN W09522610-A1.
 XX
 PD 24-AUG-1995.
 XX
 PF 15-FEB-1995; 95WO-US02644.
 XX

PR 18-FEB-1994; 94US-0199776.
 XX
 PA (HGHM) BRIGHAM & WOMENS HOSPITAL.
 XX
 PI Brenner MB, Parker CM;
 XX
 DR WPI: 1995-302716/39.
 XX
 PT New isolated integrin alpha sub-unit and peptide(s) - exhibit cell
 PT attachment activity or block activity of intra-epithelial
 PT Lymphocytes
 XX
 PS Claim 1; Page 59; 75pp; English.
 XX
 CC Peptides (AAR82638-R82655) based on, or derived from, a highly-charged
 CC fragment (AAR82636) of the X-domain of the human mucosal lymphocyte-1
 CC alpha-E chain (AAR82656) can be attached to a support to provide a
 CC surface that exhibits cell attachment activity for use e.g. in a
 CC prosthetic device or as an affinity matrix for isolating a ligand of the
 CC alpha-E chain, or can be used in assays for cpds. useful e.g. in
 CC modulating the immune system, tumour progression or metastasis,
 CC or in regulating developmental processes.
 XX
 SQ Sequence 14 AA;
 Query Match 2.1%; Score 6; DB 16; Length 14;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 218 EDEEEE 223
 DB 8 EDEEEE 13
 RESULT 28
 AAR82650
 ID AAR82650 standard; Peptide; 15 AA.
 XX
 AC AAR82650;
 XX
 DI 14-AUG-2002 (first entry)
 XX
 DE Ubiquitin binding enzyme (UBE2G1) antigenic peptide.
 XX
 KW Ubiquitin binding enzyme; UBE2G1; skeletal muscle disorder; antigen;
 KW renal glomerular sclerotic model; transgenic; enzyme; drug development.
 XX
 OS Mus sp.
 XX
 PN JP2002119170-A.
 XX
 PD 23-APR-2002.
 XX
 PF 18-OCT-2000; 2000JP-0318271.
 XX
 PR 18-OCT-2000; 2000JP-0318271.
 XX
 PA (SAKA) OTSUKA PHARM CO LTD.
 XX
 DR WPI: 2002-475324/51.
 XX
 PT Ubiquitin binding enzyme (UBE2G1) gene detected non-human animals for
 PT elucidation of pathogenic conditions of skeletal muscle related
 PT diseases and development of treatment including new drugs
 XX
 PS Example 6; Page 11; 23pp; Japanese.
 XX
 CC This invention relates to model animals without the UBE2G1 gene. These
 CC non-human animals are made substantially devoid of ubiquitin (Ub)
 CC binding enzyme activity by wholly or partially defecting or modifying
 CC the UBE2G1 gene, with other enzymes by insertion or replacement. The
 CC UBE2G1 gene is preferably substituted by a neomycin resistant gene,
 CC especially in homo-or hetero-defected type, particularly devoid of the

CC 2nd exon, and providing renal glomerular sclerotic model animals,
 CC especially mouse. The transgenic animals of the invention may be used
 CC in the development of treatments including new drugs for skeletal muscle
 CC related diseases. The present sequence represents a ubiquitin binding
 CC enzyme protein (UBE2G1) antigenic peptide used in the method of the
 CC invention.

XX Sequence 15 AA;
 SQ

Query Match 2.1%; Score 6; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 SPANVD 143
 III II
 DB 10 SPANVD 15

RESULT 29
 ABB76996
 ID ABB76996 standard; Peptide: 15 AA.
 AC ABB76996;
 XX
 XX ABB76996;
 DT 26-JUL-2002 (first entry)
 XX
 DE Human DRP-1 9.35 peptide fragment.
 XX
 KW Human: dihydropyrimidinase associated protein-1 9.35; DRP-1;
 KW nervous disease; developmental disorder.
 XX
 OS Homo sapiens.
 XX
 PN CN1331332-A.
 XX
 PD 16-JAN-2002.
 XX
 XX 26-JUN-2000; 2000CN-0116772.
 XX
 PR 26-JUN-2000; 2000CN-0116772.
 XX
 PA (BODR-) HODR GENE DEV CO LTD SHANGHAI.
 XX
 PI Mao Y. Xie Y;
 XX
 OR WPI: 2002 340575/38.
 XX
 XX A human dihydropyrimidinase associated protein-1 (DRP-1) 9.35
 PT polypeptide, and the polynucleotide encoding it, for treating e.g.
 PT nervous disease and developmental disorders -
 XX
 PS Example 5; Page 20 (Disclosure); 33pp; Chinese.
 XX
 CC The present invention relates to human dihydropyrimidinase associated
 CC protein-1 (DRP-1) 9.35 (ABB76996). DRP-1 9.35 and its coding sequence are
 CC useful for treating diseases e.g. nervous disease and developmental
 CC disorders. The present sequence is an N-terminal peptide fragment of
 CC DRP-1 9.35, which was used in an example from the invention.
 XX
 SQ Sequence 15 AA;
 Query Match 2.1%; Score 6; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 SLLHPP 100
 III I
 DB 7 SLLHPP 12

RESULT 30
 AAR82651
 ID AAR82651 standard; Peptide: 20 AA.

XX
 AC AAR82651;
 XX
 DT 25-FEB-1996 (first entry)
 XX
 DE HMI-1 alpha-E chain X domain peptide analogue.
 XX
 KW Human mucosal lymphocyte-1 antigen; HMI-1; integrin;
 KW cell attachment; T-cell; X domain.
 OS Homo sapiens.
 XX
 PN W09522610-A1.
 XX
 PD 24-AUG-1995.
 XX
 PF 15-FEB-1995; 95WO-US02044.
 XX
 PR 18-FEB-1994; 94US-0199776.
 XX
 PA (BQHM) BRIGHAM & WOMENS HOSPITAL.
 XX
 PI Brenner MB, Parker CM;
 XX
 DR WPI: 1995-302716/39.
 XX
 XX New isolated integrin alpha sub-unit and peptide(s) - exhibit cell
 PT attachment activity or block activity of intra-epithelial
 PT lymphocytes
 XX
 PS Claim 1; Page 60; 75pp; English.
 XX
 CC Peptides (AAR82638-R82655) based on, or derived from, a highly-charged
 CC fragment (AAR82636) of the X-domain of the human mucosal lymphocyte-1
 CC alpha-E chain (AAR82656) can be attached to a support to provide a
 CC surface that exhibits cell attachment activity for use e.g. in a
 CC prosthetic device or as an affinity matrix for isolating a ligand of the
 CC alpha-E chain, or can be used in assays for cpds. useful e.g. in
 CC modulating the immune system, tumour progression or metastasis,
 CC or in regulating developmental processes.
 XX
 SQ Sequence 20 AA;
 Query Match 2.1%; Score 6; DB 16; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 EDEEEE 223
 IIIII
 DB 14 EDEEEE 19

RESULT 31
 AAY78976
 ID AAY78976 standard; protein; 20 AA.
 XX
 AC AAY78976;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Canine anionic trypsinogen immunogenic fragment #1 amino acid sequence.
 XX
 KW Anionic trypsinogen; dog; monoclonal antibody production; defect;
 KW diagnose; pancreatitis; pancreatic cancer; renal insufficiency;
 KW extrapancreatic hyposecretion; immunogenic fragment.
 XX
 OS Canis familiaris.
 XX
 PN W0200009739-A1.
 XX
 PD 24-FEB-2000.
 XX
 PR 09-AUG-1999; 99WO-JP04299.

XX 10-AUG-1998; 98JP-0236604.
 PR 10-MAR-1999; 99CP-0063990.
 XX (FUJY) FUJI YAKUJIN KOGYO KK.
 XX Waritani T, Ashida Y, Yamada T;
 PI WPI; 2000-206018/16.
 DR
 XX Anti-canine trypsin monoclonal antibody, useful as reagent for quick
 PT and accurate detection and quantitation of trypsin and/or trypsin-like
 PT immunoreactants in various forms in diagnosis e.g. of pancreatic
 PT diseases
 XX
 PS Example 2: Page 33; 67pp; Japanese.
 XX This sequence represents an immunogenic fragment of the canine anionic
 CC trypsinogen amino acid sequence. The fragment corresponds to amino acid
 CC residues 139 to 158 of the anionic trypsinogen protein sequence
 CC (see AAR78975). The invention relates to monoclonal antibodies with
 CC specificity against canine trypsin, or canine trypsin-related
 CC substances. The antibodies are highly specific and can be used as reagent
 CC for quick and accurate detection and quantitation of canine trypsin and
 CC canine trypsin-like immunoreactants in various forms. The antibodies can
 CC be used in the diagnosis of diseases such as pancreatitis, pancreatic
 CC cancer, renal insufficiency and extrapancreatic hyposecretion.
 XX
 SQ Sequence 20 AA;
 Query Match 2.1%; Score 6; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 273 SGWGN 278
 DB
 AC 4 SGWGN 5
 RESULT 32
 AAR85587
 ID AAR85587 standard; Peptide; 20 AA.
 AC AAR85587;
 XX 21-MAY-2002 (first entry)
 XX Lung tumour protein L548S peptide #22.
 XX Lung tumour; cancer; T cell; immune response stimulator;
 KW cytosolic.
 XX Homo sapiens.
 XX W0200234514-A2.
 XX 17-JAN-2002.
 XX 10-JUL-2001; 2001WO-0522058.
 XX 11-JUL-2000; 2000US-0614124.
 PR 29-AUG-2000; 2000US-0651363.
 PR 08-SEP-2000; 2000US-0658824.
 PR 26-SEP-2000; 2000US-0671325.
 PR 06-OCT-2000; 2000US-0677419.
 PR 30-OCT-2000; 2000US-0702705.
 PR 13-DEC-2000; 2000US-0736457.
 PR 03-MAY-2001; 2001US-0849626.
 XX (CORI-) CORIXA CORP.
 XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
 PI Marnerakis M, Carter B, Fanger GK, Vedvick JS, Bangur CS;

PI Menabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
 XX WP; 2002-164634/21.
 XX Novel polynucleotide encoding a lung tumour polypeptide useful for
 PT stimulating and/or expanding T cells specific for a tumour protein
 XX Claim 2: SEQ ID No 1855; 223pp; English.
 XX The invention describes an isolated polynucleotide and polypeptide
 CC useful for stimulating and/or expanding T cells specific for a tumour
 CC protein for determining the presence of a cancer in a patient. A
 CC composition containing the polynucleotide and/or polypeptide is useful
 CC for treating a lung cancer in a patient. The polypeptide is useful for
 CC removing tumour cells from a biological sample. The polynucleotide is
 CC also useful as probe or primer to detect the level of mRNA encoding a
 CC tumour protein. This is the amino acid sequence of a lung tumour
 CC associated peptide, described in the method of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 20 AA;
 Query Match 2.1%; Score 6; DB 23; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 218 KDEEE 223
 DB
 AC 15 KDEEE 20
 RESULT 33
 AAR87264
 ID AAR87264 standard; peptide; 21 AA.
 AC AAR87264;
 XX 16-MAY-1996 (first entry)
 XX KEKE-motif contg. peptide used to mark other peptides for presentation.
 DE Proteasome; activation: cell-mediated immunity; immunogen; tolerance;
 KW KEKE motif.
 XX Homo sapiens.
 OS W09527058-A1.
 PN 12-OCT-1995.
 PD 01-APR-1994; 94WO-US03591.
 PF 01-APR-1994; 94WO-US03591.
 PR (UTAH) UNIV UTAH.
 PA Realini CA, Rechsteiner MC;
 XX WPI; 1995-358633/46.
 XX DNA encoding human proteasome activator used to elicit
 PT cell-mediated immunity or tolerance to a selected immunogenic
 PT peptide
 XX Disclosure; Page 38; 71pp; English.
 XX AAR87194-95, AAR87260, AAR87262-64 and AAR87266 are KEKE motif-contg.
 CC peptides which can be used in a method for inducing cell-mediated
 CC immunity against or tolerance to specific epitopes using plasmids
 CC encoding a human proteasome activator (PA) and appropriate epitope-

CC bearing peptides adjacent to presentation marker peptides (lysine and
 CC glutamine rich peptides termed KERE motifs peptides, that mark
 CC adjacent peptides for presentation). The method can be used to elicit
 CC cell-mediated immunity or tolerance to pathogen-encoded peptides or
 CC tumour specific antigens.

XX Sequence 21 AA;

Query Match 2.1%; Score 6; DB 16; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 KAEAEK 178

DB B KAEAEK 13

RESULT 34

ID AAR82636 standard; Peptide; 21 AA.

AC AAR82636;

XX

XX 25-FEB-1996 (first entry)

DE HMC-1 alpha-E chain X domain cleavage product.

XX

KW Human mucosal lymphocyte-1 antigen: HMC-1; integrin;

KW cell attachment; T-cell; X domain.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT Region 3...29

FT /label Highly-charged_region

XX

XX W09522610-A1.

PN 24-AUG-1995.

XX

XX 15-FEB-1995; 95W0-US02044.

XX

XX 18-FEB-1994; 94US-0199776.

XX

PA (BCHM) HIRSHAM & WOMENS HOSPI-AL.

XX

PI Brenner MB, Parker CM;

XX

XX WPI; 1995-302716/39.

XX

XX New isolated integrin alpha sub-unit and peptide(s) - exhibit cell
 attachment activity or block activity of intra-epithelial
 lymphocytes

XX

XX Claim 1; Page 53; 75pp; English.

XX

CC The human mucosal lymphocyte-1 alpha-E chain (AAR82636) contains an
 "X domain" (AAR82635) unique to integrins, that is believed to be
 involved in the adhesion of T-cells to epithelial cells.
 CC cleavage of the X domain yields a peptide fragment (AAR82636) that
 includes the highly-charged portion of the X domain. Peptides
 CC (AAR82638-R82655) based on this fragment show cell attachment activity.

XX

QY Sequence 21 AA;

DB 218 EDEEE 223

DB 14 EDEEE 19

RESULT 35
 AAW13209
 ID AAW13209 standard; Protein; 25 AA.

XX

AC AAW13209;

XX

DT 19-MAY-1997 (first entry)

XX

DE Fragment of p53 binding protein p53UBC.

XX

KW p53; binding protein; p53UBC; antibody; diagnosis; treatment;
 KW neoplastic; pre-neoplastic; disease; agonist; antagonist;
 KW augmentation; inhibition; complex formation; neoplasia; apoptosis;
 KW reperfusion; injury; myocardial infarction; stroke; AIDS;
 KW traumatic brain; neurodegenerative; aging; ischaemia; toxemia;
 KW infection; hepatitis; probe; genetic; forensic identification;
 KW fragment.

XX

OS Homo sapiens.

XX

PN W09514777-A1.

XX

PD 01-JUN-1995.

XX

PF 21-NOV-1994; 94W0-US13499.

XX

PR 22-NOV-1993; 93US-0156571.

XX

PA (ONYX-) ONYX PHARM.

XX

PI Bischoff JR, Wu L;

XX

DR WPI; 1995-206934/27.

XX

PI New p53-binding polypeptide(s) WBP1 and p53UBC - used to develop
 prods. for screening assays and for use in diagnosis and therapy of
 diseases, esp. neoplasia

XX

PS Claim 22; Page 41; 90pp; English.

XX

CC The present sequence is a fragment of the p53 binding protein
 p53UBC, which can be used to generate antibodies for the diagnosis of
 (pre)neoplastic diseases, and WBP1 (ant)agonists. The (ant)agonists
 CC augment or inhibit the formation of p53:p53UBC complexes and
 CC inhibit neoplasia or apoptosis, useful in the treatment of, e.g.
 CC reperfusion injury, myocardial infarction, stroke, traumatic brain
 CC injury, neurodegenerative diseases, aging, ischaemia, toxemia,
 CC infection, AIDS and hepatitis. The p53UBC encoding cDNA, which was
 CC isolated from a HeLa cell derived cDNA library using the yeast
 CC two-hybrid system, can be used to generate probes for the diagnosis
 CC of (pre)neoplastic pathological conditions and genetic diseases,
 CC and the forensic identification of human individuals.

XX

SQ Sequence 25 AA;

Query Match 2.1%; Score 6; DB 16; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 LLNEPN 135

DB 3 LLNEPN 8

RESULT 36

AAW57374
 ID AAW57374 standard; peptide; 25 AA.

XX

AC AAW57374;

XX

DT 11-AUG-1998 (first entry)

XX

DE Human p53UBC immunogenic peptide SEQ ID NO:108.
 XX
 KW Human; WBPI; p53; cancer; interacting protein; screen; diagnosis;
 KW genetic disease; forensic identification; nuclear phosphoprotein;
 KW cellular proliferation; neoplastic transformation; p53UBC; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN US5756649-A.
 XX
 PU 26-MAY-1998.
 XX
 PF 02-MAR-1995; 95US-0399696.
 XX
 PR 02-MAR-1995; 95US-0399696.
 PR 22-NOV-1993; 93US-0156571.
 PR 21-NOV-1994; 94WO-US13499.
 XX
 PA (ONYX-) ONYX PHARM INC.
 XX
 PI Bitschoff JR, W3 L3;
 XX
 DR WPI: 1998-321626/28.
 XX
 PT WBPI, p53 interacting protein - useful screening agents for
 PT treatments of p53 related cancers
 XX
 PS Disclosure; Column 27; 68pp; English.
 XX
 CC The present sequence represents an immunogenic peptide from p53UBC,
 CC from the present invention describing WBPI (cellular proliferation and
 CC neoplastic transformation) polypeptide, which can be used to screen
 CC bacteriophage antibody display libraries or to immunize a rabbit. WBPI
 CC is a p53-interacting protein that can be used to screen for agents for
 CC diagnosis or treatment of cancer and genetic diseases and for forensic
 CC identification of human individuals. The p53 protein is a nuclear
 CC phosphoprotein involved in control of cellular proliferation, and
 CC mutations in it are associated with human cancers.
 XX
 SQ Sequence 25 AA:
 Query Match 2.1%; Score 6; DB 19; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 130 LANEPN 135
 DB 3 LANEPN 8
 RESULT 37
 AAB30911
 ID AAB30911 standard; peptide: 25 AA.
 XX
 AC AAB30911:
 XX
 DI 02-APR-2001 (first entry)
 XX
 DE Amino acid sequence of peptide linking fluorophores in FRET-reporters.
 XX
 KW Apoptosis reporter; fluorescence resonance energy transfer; FRET;
 KW fluorophore; apoptosis; apoptosis inducer; cell survival.
 XX
 OS Unidentified.
 XX
 PN W0200075160-A1.
 XX
 PU 14-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US15142.
 XX
 PR 04-JUN-1999; 99US-0326472.
 XX

PA (MERI) MERCK & CO INC.
 XX
 PI Elliott KJ, Kounnas MZ, Dyer KJ, Munoz R, Wagner SL, Jones JM;
 PI Corey-Naeve J;
 XX
 DR WPI: 2001-091198/10.
 DR N-PSDB; AAC68837.
 XX
 PI New recombinant apoptosis reporter cells expressing fluorescence
 PI resonance energy transfer reporter polypeptides, useful in screening
 PI assays for identifying and selecting compounds that modulate apoptosis
 XX
 PS Disclosure; Fig 1A; 68pp; English.
 XX
 CC The specification describes a stable recombinant apoptosis reporter
 CC cell line which expresses a fluorescence resonance energy transfer
 CC (FRET) apoptosis reporter polypeptide. This polypeptide comprises an
 CC exciter fluorophore peptide bonded with a linker polypeptide, that is
 CC in turn peptide bonded with an emitter fluorophore. The cell line is
 CC used to detect apoptosis induction by an apoptosis inducer agent. The
 CC recombinant apoptosis reporter cells are used in screening assays for
 CC identifying and selecting compounds that modulate apoptosis. They are
 CC useful for screening unknown compounds to identify those which modulate
 CC apoptosis induced by an extracellular apoptosis inducer agent, or
 CC apoptosis induced by removal of a critical factor for cell survival.
 CC The present sequence represents a peptide linking fluorophores in
 CC FRET-reporters of the invention.
 XX
 SQ Sequence 25 AA:
 Query Match 2.1%; Score 6; DB 22; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 DLYDDD 214
 DB 12 DLYDDD 17
 RESULT 38
 AAR54953
 ID AAR54953 standard; Protein: 26 AA.
 XX
 AC AAR54953;
 XX
 DI 31-OCT-1994 (first entry)
 XX
 DE Agraal grass pollen allergen allergenic determinant.
 XX
 KW IgE; allergy; antigen; diagnosis; treatment; Group I allergen.
 XX
 OS Pooidae.
 XX
 PN W09410314-A.
 XX
 PU 11-MAY-1994.
 XX
 PF 29-OCT-1993; 93WO-A000559.
 XX
 PR 30-OCT-1992; 92US-0971096.
 XX
 PA (UYME) UNIV MELBOURNE.
 XX
 PI AviJooglu A, Knox RB, Singh MB;
 XX
 DR WPI: 1994-167469/20.
 XX
 PT DNA encoding allergenic proteins and peptide(s) from Johnson
 PT grass pollen allergen Sorhl - for diagnosing, treating and
 PT preventing allergy to Johnson pollen
 XX
 PS Disclosure; Page 53; 81pp; English.

```

XX  the sequence is that of an Agral Group 1 grass pollen allergen internal
CC  allergenic determinant. The peptide sequence was compared to
CC  allergenic determinant sequences of four other Group 1 allergens
CC  (all members of the subfamily Poideae) and homology used to determine
CC  the antigenic determinants of Sorhl.
CC  See also AAR54949-76.
XX
XX
SQ  Sequence 26 AA:

Query Match      2.1%; Score 6; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  172 TKAEAE 177
DB  7 TKAEAE 12

RESULT 39
AAR54354
ID  AAR54954 standard; Protein: 26 AA.
XX
AC  AAR54954:
XX
DT  31-OCT-1994 (first entry)
XX
DE  PoaI grass pollen allergen allergenic determinant.
XX
KW  IgE; allergy; antigen; diagnosis; treatment; Group 1 allergen.
XX
CS  Poideae.
XX
PN  W03410314-A.
XX
PD  11-MAY-1994.
XX
PF  29-OCT-1993; 93WO-AU00559.
XX
PR  30-OCT-1992; 92US-C971056.
XX
PA  {UYME } UNIV MESSBOURNE.
XX
PI  Avijooğlu A, Knox RB, Singh MB;
XX
DR  WPI: 1994-167469/20.
XX
CC  DNA encoding allergenic proteins and peptide(s) from Johnson
CC  grass pollen allergen Sorhl - for diagnosing, treating and
CC  preventing allergy to Johnson pollen.
XX
PS  Disclosure; Page 54; 81pp; English.
XX
CC  The sequence is that of a PoaI Group 1 grass pollen allergen internal
CC  allergenic determinant. The peptide sequence was compared to
CC  allergenic determinant sequences of four other Group 1 allergens
CC  (all members of the subfamily Poideae) and homology used to determine
CC  the antigenic determinants of Sorhl.
CC  See also AAR54949-76.
XX
XX
SQ  Sequence 26 AA:

Query Match      2.1%; Score 6; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  172 TKAEAE 177
DB  7 TKAEAE 12

RESULT 40
AAR58723

```

```

ID  AAY88723 standard; peptide; 28 AA.
XX
AC  AAY88723:
XX
DT  23-MAY-2000 (first entry)
XX
DE  Core polypeptide fragment I No. 78.
XX
KW  Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;
KW  HIV-2; SIV; pharmacokinetic; half-life; growth factor; cytokine; viral;
KW  anti-fusogenic; differentiation factor; interleukin; interferon;
KW  colony stimulating factor; hormone; angiogenic factor.
XX
CS  Unidentified.
XX
PN  W0959615-A1.
XX
PD  25-NOV-1999.
XX
PF  20-MAY-1999; 99WO-US11219.
XX
PR  20-MAY-1998; 98US-0082279.
XX
PA  (TRIM-) TRIMERIS INC.
XX
PI  Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
XX
DR  WPI: 2000-136792/12.
XX
CC  A new hybrid polypeptide with enhanced pharmacokinetic properties
CC  comprises enhancer sequence -
XX
PS  Disclosure; Page 22; 124pp; English.
XX
CC  The invention relates to hybrid polypeptides comprising enhancer peptide
CC  sequence linked to core polypeptides. The enhancer polypeptides are
CC  derived from various retroviral envelope (gp41) protein sequences,
CC  especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the
CC  pharmacokinetic properties such as increasing the half-life of any core
CC  polypeptide that they are linked to. The core polypeptides are any
CC  polypeptide that may be introduced into a living system and that can
CC  function as a pharmacologically useful peptide for the treatment or
CC  prevention of a disease. The core polypeptides are bioactive peptides
CC  selected from a growth factor, cytokine, differentiation factor,
CC  interleukin, interferon, colony stimulating factor, hormone or
CC  angiogenic factor. The peptides of the invention can be used for
CC  inhibiting viral infection and can be used in anti-viral and
CC  anti-fusogenic treatments. Sequences AAY88651-Y90055 represent core
CC  polypeptide fragments that can be used in the invention. Some sequences
CC  among those indicated also comprise enhancer fragments at terminal ends
CC  and form hybrid polypeptides.
XX
SQ  Sequence 28 AA:

Query Match      2.1%; Score 6; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  218 EDEEEE 223
DB  18 EDEEFF 23

Search completed: April 10, 2003, 10:39:19
Job time : 33.5271 secs

```



```

; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/976,165
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/565,538
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/055,699
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/820,170
; PRIOR FILING DATE: 1997-03-19
; PRIOR APPLICATION NUMBER: JP 63410/1996
; PRIOR FILING DATE: 1996-03-19
; PRIOR APPLICATION NUMBER: JP 69163/1997
; PRIOR FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patent; Ver. 2.1
; SEQ ID NO 22
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo sapiens
US 09-976-165 22

Query Match 39.8%; Score 460.5; DB 10; Length 170;
Best Local Similarity 52.4%; Pred. No. 2.9e-36;
Matches 87; Conservative 28; Mismatches 44; Indels 7; Gaps 3;

QY 6 MISSOAKA:MI---EKSQLEQEVGCFRTILVDSDLYNWEVA:FGLPN:LYEGYFKAHTK 63
DB 1 M:EGSALH:RQAALAKNKNIVEGFSAGLIDND:YRWEV:GILGPDILYEGGVFKAHIT 60
QY 64 EP:DYPSPPTFRELTKMWHNIYENSDVCISILHIPPVDDPQSGELPSEKWNPTQNVRTI 123
DB 61 FPKGVP:LRPKKFKITETWHPNDKNDVCISILHPEQSDKYGVKPKFERNWLIHTVETI 120
QY 124 LLSVLS:LNPNIESPANVDASVMFRKRDGSKGK:KEYAE:IRK 167
DB 121 MSLVSL:ADPNDS:SPANVDAA---KEMREDRNGEFGKRVKVRK 163

RESULT 4
US-10-060-013-22
; Sequence 22, Application US/10/60019
; Publication No. US2003003564A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Mike
; APPLICANT: Williams, Andrew
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
; TITLE OF INVENTION: DEPENDENT PROTEOLYSIS
; FILE REFERENCE: 11757.10US01
; CURRENT APPLICATION NUMBER: US/10/060,019
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US/09/177,165
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/092,443
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/063,254
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent; Ver. 2.1
; SEQ ID NO 22
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-060-013-22

Query Match 28.6%; Score 432.5; DB 9; Length 295;
Best Local Similarity 33.3%; Pred. No. 2.9e-33;
Matches 93; Conservative 61; Mismatches 78; Indels 47; Gaps 8;

QY 1 MAQUQNTSSQKALM:FIKSIQF---EPVEGFRTILVDSDLYNWEVA:FGLPN:LYEGGY 57

```

```

DB 1 MSSKSIAS-SLLROYRELDIPKKAIPSPHIEEDDSNIFTWNGVMV:NEDS:YHOGF 59
QY 58 FKAHKEP:DYPSPPTFRELTKMWHNIYENGVDVCISILHIPPVDDPQSGELPSEKWNPT 117
DB 60 FKAOMREPDPFSPPOKFTPAIYHNVDGRCICISILHQS-GDPMTDPPDAFTWSPV 118
QY 118 QNVRTILLSVLS:LNPNIESPANVDASVMFRKRDGSKGK:KEYAE:IRKQVSA:KAEAF 177
DB 119 QTVESVLISVLSLEDPNINS:SPANVDAVDYRK-----NPEYKQRYKMEVERSKQDIP 172
QY 178 KGVKVP:FIHAECYCIKTK-----VPSNDNSDLYLYDLYD----- 213
DB 173 K-GFIMPIS:SAVISOSKLODEPESKNOMADNFWYDSLDODDENGSV:ILQDIDYDDGNH 231
QY 214 -----DIDDEFEEDACDYDDEDSGMRSDVL 241
DB 232 PFEDDDVYNYNNDODHRIEFEDDDDDDDDDSDIDSDSV 270

RESULT 5
US-09-842-528-4
; Sequence 4, Application US/09842528
; Patent No. US2002004236A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: 27960, A NOVEL UBIQUITIN-CONJUGATING
; TITLE OF INVENTION: ENZYME FAMILY MEMBER AND USES THEREFOR
; FILE REFERENCE: 10448-045001
; CURRENT APPLICATION NUMBER: US/09/842,528
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-09-842-528-4

Query Match 27.4%; Score 413.5; DB 10; Length 165;
Best Local Similarity 48.8%; Pred. No. 8.3e-32;
Matches 79; Conservative 28; Mismatches 54; Indels 1; Gaps 1;

QY 7 TSSQKALM:FIKSIQF---EPVEGFRTILVDSDLYNWEVA:FGLPN:LYEGYFKAH:KFP 66
DB 4 SSASKRLQELKKLOEDPPGFSAGPVDDNNIFENNVILMGPPDTPYEGGVK:LEIEFPE 63
QY 67 DYPSPPTFRELTKMWHNIYENSDVCISILHIPPVDDPQSGELPSEKWNPTQNVRTI 125
DB 64 DYPKPKV:RFTSKIFHPNIYSNTGRTCLD:ILKDPGDDPNQYFAEENKWS:PAVTFVSILL 123
QY 126 SVLS:LNPNIESPANVDASVMFRKRDGSKGK:KEYAE:IRK 167
DB 124 SIQSLSDPNPDS:SPANVDAVDYRKEDREYKRVKWEV 165

RESULT 6
US-10-157-669-17
; Sequence 17, Application US/10157669
; Publication No. US2003005495A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Jackson, Jennifer L.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN UBIQUITIN-CONJUGATING ENZYMES
; FILE REFERENCE: PF-0356-2 CIP
; CURRENT APPLICATION NUMBER: US/10/157,669
; CURRENT FILING DATE: 2002-05-28

```

; PRIOR APPLICATION NUMBER: 09/520,076; 09/359,967; 08/965,689; 08/933,750
 ; PRIOR FILING DATE: 2000-03-07; 1999-07-22; 1997-11-06; 1997-09-23
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 17
 ; LENGTH: 165
 ; TYPE: PRI
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Accyle ID No. US20030054385A1 2456290CBI
 US-10-157-669-17

Query Match 25.2%; Score 380.5; DB 9; Length 165;
 Best Local Similarity 50.3%; Pred. No. 1.1e-28;
 Matches 80; Conservative 1d; Mismatches 54; Indels 7; Gaps 2;

Qy 11 KALMLELSKIQHPVEGFRITLVDSOLYNWEVAIFGLPNTLYEGGYFKAHKEFDY 70
 Db 7 KRLMAEYKQLTINPEGVACPMKNEFFWEALIMGPDTCFEGVFAILSFIDYPL 66
 Qy 7: SPETNELTKMMHNTIYENGUCVISTILHPPVDSQSGELPSEKRNPTQNVRTILSVLSL 130
 Db 67 SPKMRFCCEMHPNITPDGRVCLSIHAPGDDPMGYESSAPRSPVQSEKTLISVWSM 126
 Qy 13: LNEPNTFSPANVDASVMEKRWDSKQKYEAFILRKQV 169
 Db 127 LAEPNDSGANYDASKM WRD---DREQFYKIAQI 158

RESULT 7
 US-10-157-669-33
 ; Sequence 33, Application US/17157669
 ; Publication No. US20030054385A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gal, Preeti G.
 ; APPLICANT: Jackson, Jennifer L.
 ; APPLICANT: Corley, Neil C.
 ; TITLE OF INVENTION: HUMAN UBIQUITIN-CONJUGATING ENZYMES
 ; FILE REFERENCE: PF-0356.2 CIP
 ; CURRENT APPLICATION NUMBER: US/10/157,669
 ; CURRENT FILING DATE: 2002-05-28
 ; PRIOR APPLICATION NUMBER: 09/520,076; 09/359,967; 08/965,689; 08/933,750
 ; PRIOR FILING DATE: 2000-03-07; 1999-07-22; 1997-11-06; 1997-09-23
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 33
 ; LENGTH: 165
 ; TYPE: PRI
 ; ORGANISM: Saccharomyces cerevisiae
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: Genbank ID No. US20030054385A1 g4257
 US-10-157-669-33

Query Match 23.7%; Score 358.5; DB 9; Length 165;
 Best Local Similarity 47.3%; Pred. No. 1.4e-26;
 Matches 70; Conservative 26; Mismatches 49; Indels 3; Gaps 1;

Qy 8 SSOKALMLELSKIQHPVEGFRITLVDSOLYNWEVAIFGLPNTLYEGGYFKAHKEFD 67
 Db 4 TAOKRLKALQGLIKDSPPGCVAGPSKSNFIWDCILQGPPTPYADGVFNKAEFPKD 63
 Qy 68 YPSHPPTFRELTKMMHNTIYENGUCVISTILHPPVDSQSGELPSEKRNPTQNVRTILSV 127
 Db 64 YPLSPKLTFTPSLIHPNIYNGEVCISILHSPGDDPMNYELAEERSPVQSEKILLSV 123
 Qy 128 ISLLENPNTFSPANVDASVMEKRWDSK 155
 Db 124 MSMLSEPNIESGANTDAICL---WRQNR 148

RESULT 8
 US-09-942-528-5

; Sequence 5, Application US/09842528
 ; Patent No. US20020004236A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Meyers, Rachel A.
 ; APPLICANT: Tsai, Fong-Ying
 ; TITLE OF INVENTION: 27960, A NOVEL UBIQUITIN CONJUGATING
 ; TITLE OF INVENTION: ENZYME FAMILY MEMBER AND USES THEREFOR
 ; FILE REFERENCE: 10448-045001
 ; CURRENT APPLICATION NUMBER: US/09/842,528
 ; CURRENT FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: US 60/199,500
 ; PRIOR FILING DATE: 2000-04-25
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSQ for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 144
 ; TYPE: PRI
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: consensus sequence
 US-09-842-528-5

Query Match 21.9%; Score 331.5; DB 10; Length 144;
 Best Local Similarity 43.4%; Pred. No. 4.3e-24;
 Matches 69; Conservative 27; Mismatches 42; Indels 21; Gaps 4;

Qy 11 KALMLELSKIQHPVEGFRITLVDSOLYNWEVAIFGLPNTLYEGGYFKAHKEFDY 70
 Db 1 KRLKELKELQKDPGP--ISAIPVDNLEWTVIVGPDTPYEGGVFKLTIEFPEYFF 58
 Qy 71 SPPHFLTKMMHNTIYENGUCVISTILHPPVDSQSGELPSEKRNPTQNVRTILSVIS 129
 Db 59 KPPKVFITKIYHPNVDSGSEICLDIL-----KEKWSPALITLTVLLSIQS 105
 Qy 130 LNEPNTFSPANVDASVMEKRWDSKQK-----DKFYAE 163
 Db 106 LNEPNDPSPCNVDAAELYKRDREPKKKVREWTKKYAE 144

RESULT 9
 US-09-925-301-866
 ; Sequence 868, Application US/09925301
 ; Patent No. US20020052308A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA106
 ; CURRENT APPLICATION NUMBER: US/09/925,301
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05882
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 1694
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 868
 ; LENGTH: 196
 ; TYPE: PRI
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (6)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (9)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-925-301-868

Query Match 19.7%; Score 297; DB 10; Length 196;
 Best Local Similarity 37.1%; Pred. No. 1.2e-20;
 Matches 66; Conservative 31; Mismatches 57; Indels 24; Gaps 4;

Qy 2 AQQQMTSSQKALMLELSKIQHPVEGFRITLVDSOLYNWEVAIFGLPNTLYEGGYFKAH 61


```

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ For Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-785-671-2

Query Match      15.6%   Score 235.5; DB 10; Length 158;
Best Local Similarity 35.6%   Pred. No. 6.5e-15;
Matches 57; Conservative 21; Mismatches 61; Indels 21; Gaps 3;

DB 42 AALMSDARRKLRKFRKQCEQDPVGVGVS-CAPSENNIMUNNAVIFGPGFEDGTFKLV 100
QY 62 KEPIIDVPYSPTEFLTKMHPNIYENGDCVTSILHPPVDDPQSGELPSERWNPNTQVR 121
DB 101 IEFSEYYPNKPVTFRFSKMHHPNYADGSCICDIL-----QNRWSPYDVS 147
QY 122 TLLSVISLILNPNTSPANVDASVMFRKWRKSKGKKEYAFIRKUVSAIKAPAKKD 179
DB 148 SILTSIQSLLEPNPNPNPANQAALYOE-----NKREY-----EKRVAIVEQSND 195

RESULT 10
US-10-108-605-331
; Sequence 331, Application US/101C8695
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 311338
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 03/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 36;
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 331
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-331

Query Match      15.7%   Score 237; DB 9; Length 147;
Best Local Similarity 31.8%   Pred. No. 4.2e-15;
Matches 53; Conservative 31; Mismatches 56; Indels 20; Gaps 3;

QY 11 KALMLEKSLQEPVEGFRITLVDESD---LYNNEVAIFGLPNTLYEGGYKAKHIKFPIDY 70
DB 4 KRINKELQGLRFPAGQAGVGD-DLFHQWQIMGPPDSYQGGVFLLIHFPIDYFP 62
QY 71 SPPTFRFLTKMHPNIYENGDCVTSILHPPVDDPQSGELPSERWNPNTQVRTILLSVIS 130
DB 63 KPKYVATFRYHPNINSNGSICLALK-----SQNSPALILSKVLLSICSL 109
QY 131 LNEPNTSPANVDASVMFRKWRKSKGKKEYAEIRK 167
DB 110 LCPNPQDPLVPEIARIYKTKRE-----KYNELARE 140

RESULT 11
US-09-785-671-2
; Sequence 2, Application US/09785671
; Patent No. US20020128189A1
; GENERAL INFORMATION:
; APPLICANT: Kho, Choon-Joo
; APPLICANT: Lee, Mu-Fu
; APPLICANT: Haber, Edgar
; TITLE OF INVENTION: UBQUITINATION OF THE TRANSCRIPTION FACTOR E2A
; FILE REFERENCE: 05433-035062
; CURRENT APPLICATION NUMBER: US/09/785,671
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 08/986,605
; PRIOR FILING DATE: 1997-12-08
; PRIOR APPLICATION NUMBER: US 08/825,476
; PRIOR FILING DATE: 1997-03-28
; PRIOR APPLICATION NUMBER: US 60/614,388
; PRIOR FILING DATE: 1996-03-28
```

```

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ For Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-785-671-2

Query Match      15.6%   Score 235.5; DB 10; Length 158;
Best Local Similarity 35.6%   Pred. No. 6.5e-15;
Matches 57; Conservative 21; Mismatches 61; Indels 21; Gaps 3;

QY 13 LMLEKSLQEPVEGFRITLVDESD---LYNNEVAIFGLPNTLYEGGYKAKHIKFPIDY 68
DB 9 LAQERKAWRKDHDFGFGVAVPTKNPCTMTNLMNNECAIPCKGTPWEGGLFKRLMKDDY 64
QY 69 PYSPTFRFLTKMHPNIYENGDCVTSILHPPVDDPQSGELPSERWNPNTQVRTILLSVI 128
DB 69 PSSPPCKCFEPPLFHPNVYPSGIVCLSTLEEDKD-----WRPAITIKOILIGIQ 117
QY 129 SLLNEPNTSPANVDASVMFRKWRKSKGKKEYAEIRKQ 168
DB 118 ELLNEPNIQDPAQAEAYTYCONR-----VEYEKRVRAQ 151

RESULT 12
US-09-925-302-525
; Sequence 525, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCI/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 525
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-525

Query Match      15.6%   Score 235.5; DB 10; Length 205;
Best Local Similarity 35.6%   Pred. No. 9.3e-15;
Matches 57; Conservative 21; Mismatches 61; Indels 21; Gaps 3;

QY 13 LMLEKSLQEPVEGFRITLVDESD---LYNNEVAIFGLPNTLYEGGYKAKHIKFPIDY 68
DB 56 LAQERKAWRKDHDFGFGVAVPTKNPCTMTNLMNNECAIPCKGTPWEGGLFKRLMKDDY 115
QY 69 PYSPTFRFLTKMHPNIYENGDCVTSILHPPVDDPQSGELPSERWNPNTQVRTILLSVI 128
DB 116 PSSPPCKCFEPPLFHPNVYPSGIVCLSTLEEDKD-----WRPAITIKOILIGIQ 164
QY 129 SLLNEPNTSPANVDASVMFRKWRKSKGKKEYAEIRKQ 168
DB 165 ELLNEPNIQDPAQAEAYTYCONR-----VEYEKRVRAQ 198

RESULT 13
US-10-043-487-394
; Sequence 394, Application US/0043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polype
; TITLE OF INVENTION: mammalian polypeptides
```



```
Query Match: 15.3%, Score 231.5, DB 9; Length 197;
Best Local Similarity 33.7%, Pred. No. 2, 1e-14;
Matches 55; Conservative 21; Mismatches 64; Indels 23; Gaps 4;

QY 13 LMELKSLQEEPEVEGRITLVDSGLYNNWEVAIFGLPNTLYEGGYKAHKFPIDYPS 71
DB 7 LKRELHMLAEPPPG--ITCWQXQDQADLRAQILGGANTPYEKGVTLEVIIPERYPE 64
QY 72 PPTFRLTKMHPNIYENGVCISILHPPVDDPQSGELPSEKWNPTQNVRTILLSVISL 131
DB 65 PPTFRLTKMHPNIYENGVCISILHPPVDDPQSGELPSEKWNPTQNVRTILLSVISL 115
QY 132 NEPTSPANVDASVNF-----RKWRDSKCKEYAE 163
DB 116 SEPNTFPLMADISSEFKYNTAFUKKAKQWTEKHARQKQKAD 158

RESULT 17
US-09-826-312-4
: Sequence 4, Application US/09826312
: Patent No. US2002004208A;
: GENERAL INFORMATION:
: APPLICANT: Issakani, Sarkiz D.
: APPLICANT: Huang, Jianing
: APPLICANT: Shuang, Julie
: APPLICANT: Pray, Todd R.
: TITLE OF INVENTION: UBIQUITIN LIGASE ASSAY
: FILE REFERENCE: A-68603 1/RMS/JJD
: CURRENT APPLICATION NUMBER: US/09/826,312
: CURRENT FILING DATE: 2001-04-03
: PRIOR APPLICATION NUMBER: US 09/542,497
: PRIOR FILING DATE: 2000-04-03
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: Patent in version 3.1
: SEQ ID NO 4
: LENGTH: 148
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-826-312-4

Query Match: 15.0%, Score 226; DB 10; Length 148;
Best Local Similarity 31.2%, Pred. No. 4, 8e-14;
Matches 49; Conservative 20; Mismatches 59; Indels 20; Gaps 3;

QY 13 KALMLELKS-QRPVEGRITLVDSGLYNNWEVAIFGLPNTLYEGGYKAHKFPIDY 70
DB 4 KRINKELSLDARPPAQSAGPVGD-DMFHQWATIMGPNDSFYQGVFELTIHPTDY 62
QY 71 SP2TFLTKMHPNIYENGVCISILHPPVDDPQSGELPSEKWNPTQNVRTILLSVISL 130
DB 63 KPPKVAFTIHYHPNTNSGSLDILR-----SOWSPAATISKVLLSTCSL 109
QY 131 LNEPTSPANVDASVNF-----RKWRDSKCKEYAE 167
DB 110 LCPDPLVPEIARIYKTDQ-----KYNKISRE 140

RESULT 18
US-10-180-927-2
: Sequence 2, Application US/10180927
: Publication No. US20030046721A1
: GENERAL INFORMATION:
: APPLICANT: Lovitch, Michael W.
: TITLE OF INVENTION: HSPC150-LIKE GENE DISRUPTIONS.
: FILE REFERENCE: R-437
: CURRENT APPLICATION NUMBER: US/10/180,927
: CURRENT FILING DATE: 2002-06-25
: PRIOR APPLICATION NUMBER: US 60/301,281
: PRIOR FILING DATE: 2001-06-26
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
```

```
: SEQ ID NO 2
: LENGTH: 204
: TYPE: PRT
: ORGANISM: Mus musculus
US-10-180-927-2

Query Match: 14.9%, Score 225.5; DB 9; Length 204;
Best Local Similarity 32.0%, Pred. No. 8, 2e-14;
Matches 57; Conservative 25; Mismatches 73; Indels 23; Gaps 4;

QY 13 LMELKSLQEEPEVEGRITLVDSGLYNNWEVAIFGLPNTLYEGGYKAHKFPIDYPS 71
DB 7 LKRELHMLAEPPPG--ITCWQXQDQADLRAQILGGANTPYEKGVTLEVIIPERYPE 64
QY 72 PPTFRLTKMHPNIYENGVCISILHPPVDDPQSGELPSEKWNPTQNVRTILLSVISL 131
DB 65 PPTFRLTKMHPNIYENGVCISILHPPVDDPQSGELPSEKWNPTQNVRTILLSVISL 115
QY 132 NEPTSPANVDASVNF-----RKWRDSKCKEYAEIRKQVSNATKAEAK 178
DB 116 AFPNDPDLMDITSEFKYNTAFUKKAKQWTEKHARQKQKADDEELGTSSEVGDSEF 173

RESULT 19
US-10-043-487-392
: Sequence 392, Application US/10043487
: Publication No. US20030055220A1
: GENERAL INFORMATION:
: APPLICANT: HYBRIGENICS
: APPLICANT: Pierre, LEGRAIN
: TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
: FILE REFERENCE: B4778A
: CURRENT APPLICATION NUMBER: US/10/043,487
: CURRENT FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/261,130
: PRIOR FILING DATE: 2001-01-12
: NUMBER OF SEQ ID NOS: 561
: SOFTWARE: Patent in version 3.1
: SEQ ID NO 392
: LENGTH: 143
: TYPE: PRT
: ORGANISM: Shigella Flexneri
US-10-043-487-392

Query Match: 14.8%, Score 224; DB 9; Length 143;
Best Local Similarity 31.6%, Pred. No. 7e-14;
Matches 48; Conservative 28; Mismatches 56; Indels 20; Gaps 3;

QY 16 ELKSLQEEPEVEGRITLVDSGLYNNWEVAIFGLPNTLYEGGYKAHKFPIDYPSPTTF 75
DB 5 ELSLDLARDPPAQSAGPVGD-DMFHQWATIMGPNDSFYQGVFELTIHPTDYKPKV 63
QY 76 RELTKMHPNIYENGVCISILHPPVDDPQSGELPSEKWNPTQNVRTILLSVISLNEPN 135
DB 64 AETTRIYHPNTNSGSLDILR-----SOWSPAATISKVLLSGSLUCDEN 110
QY 136 TFSADVNDASVNF-----RKWRDSKCKEYAEIRK 167
DB 111 PDDPLVPEIARIYKTDQ-----KYNKISRE 136

RESULT 20
US-10-043-487-399
: Sequence 399, Application US/10043487
: Publication No. US20030055220A1
: GENERAL INFORMATION:
: APPLICANT: HYBRIGENICS
: APPLICANT: Pierre, LEGRAIN
: TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
: FILE REFERENCE: B4778A
: CURRENT APPLICATION NUMBER: US/10/043,487
```

```

Query Match      13.58:  Score 204:  DB 10:  Length 161;
Best Local Similarity 32.78:  Pred. No. 6.6e-12:
Matches 52:  Conservative 26:  Mismatches 53:  Indels 28:  Gaps 5;

QY 6 MTSOKALMLELKSIFQFEFVEGRFRIIVDES---DLYNMFVAIFQINLYEGYGFKAHI 62
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1: MASMOKRLQKELLALNDPPGQ--MTLNKSVNSITQIVDMEGAPQIYEGEKFOLE 68

```



```

; SEQ ID NO 979
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-979

Query Match 13.9% Score 196; DB 10; Length 283;
Best Local Similarity 27.9%; Pred. No. 8.2e-11;
Matches 50; Conservative 31; Mismatches 70; Indels 28; Gaps 4;

QY 14 ELKSIOPEVGFRIILVDESILYNWEVA:FGLPNTLYEGGYFAHAKFKPIDYSPPTIF 75
DB 60 EVTITADPDGIKV-FPNEEDLTGLQVTFEGESTPYAGSELRMKLLGKQCPASPPKG 138
QY 76 RELTKMHPNIYENGDCVCSIL:HPVDVDPQSGELPSEKWNPIQNVRIILSVISLLNEPN 135
DB 139 YFUTKIFHPNVANGEL:GVNVL-----KRWTAHIGIRVILITKCLLIHN 185
QY 136 TESPANVDAS-VYFRKKWD-----SKGDKKEYAEL:KKQVSATKADEKDG 180
DB 186 PESAINFEAGRI:ILENYEYAAARALITE:HCGAGCPGSCRAEAGKALASGTASSTDPG 244

RESULT 28
US-10-157-669-32
; Sequence 32, Application US/10157669
; Publication No. US20030054385A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Jackson, Jennifer L.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN URIQUITTIN-CONJUGATING ENZYMES
; FILE REFERENCE: PF-0356-2 C18
; CURRENT APPLICATION NUMBER: US/10157-669
; CURRENT FILING DATE: 2002-05-26
; PRIOR APPLICATION NUMBER: 09/540,376; 09/359,967; 08/965,689; 08/933,750
; PRIOR FILING DATE: 2000-03-07; 1999-07-22; 1997-11-06; 1997-09-23
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL PROGRAM
; SEQ ID NO 32
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No. US20030054385A1 g1628097
US-10-157-669-32

Query Match 12.8% Score 193; DB 9; Length 180;
Best Local Similarity 31.0%; Pred. No. 8.5e-11;
Matches 49; Conservative 36; Mismatches 57; Indels 16; Gaps 6;

QY 11 KALMLRLK----SIOPEVGFRIILVDESILYNWEVA:FGLPNT-LYEGGYFAHAKFKP 65
DB 27 KLLAQELQLEALRIQKQKLIWHLEVSTSLCHELELV--TPQEGYRGKGFRTITVP 84
QY 66 IDYPSPTFRF:TKMHPNIYENGDCVCSIL:HPVDVDPQSGELPSEKWNPIQNVRIIL 125
DB 85 PEYNNVPPVVKCI:TKMHPNIYENGDCVCSIL:HPVDVDPQSGELPSEKWNPIQNVRIIL 137
QY 126 SVISLLNEPNFSPA-NVDASVYFRKKWD:SKGDKKEY 161
DB 136 GLVSLNLMDFNDALNIQAQMSQNSRESFNHVRVY 175

RESULT 29
US-10-043-487-395
; Sequence 395, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: Hybrigenics
; APPLICANT: Pierre, Ibrahim
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; TITLE OF INVENTION: mammalian polypeptides
```

```

; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 395
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Shigella Flexneri
US-10-043-487-395

Query Match 12.7% Score 192.5; DB 9; Length 153;
Best Local Similarity 30.1%; Pred. No. 7.6e-11;
Matches 46; Conservative 29; Mismatches 55; Indels 23; Gaps 5;

QY 6 MTSQKALMLEKLSQEPVGFRIILVDESILYNWEVA:FGLPNTLYEGGYFAHAKFKP 64
DB 1 MNASMR-VYKELEDLQKKPPPYLRLNLSDDANLYVHIAL:--LFDQPPYHLKAEFLRISF 57
QY 65 PIDYPSPTFRF:TKMHPNIYENGDCVCSIL:HPVDVDPQSGELPSEKWNPIQNVRIIL 124
DB 58 PPEYFKPMIRKFTTIRHPNVDENGOICLPI-----ISSENWKPCTKCCQVL 105
QY 125 LSVISLLNEPNFSPA-NVDAS-----VYFRKK 150
DB 106 FALNVLNVRNIRREPLRMDLADULLTQNPELFRK 138

RESULT 30
US-09-989-920-204
; Sequence 204, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 204
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-204

Query Match 12.2% Score 184; DB 9; Length 252;
Best Local Similarity 31.1%; Pred. No. 9.7e-10;
Matches 50; Conservative 22; Mismatches 55; Indels 34; Gaps 6;

QY 19 SIOPEVGFRIILVDESILYNWEVA:FGLPNTLYEGGYFAHAKFKPIDYSPPTIFRFL 78
DB 2 SIYKEPPPGM-FVVDVDMTKIUALITGFDTPYEGGFFLVFRCPDPDPIHPVRVKLM 60
QY 79 TK-----MHPNIYENGDCVCSIL:HPVDVDPQSGELPSEKWNPIQNVRIILSVISLL 131
DB 61 TTONTVRENPENYRNCKVCLSL-----GTLGPAWSPAQSISSVL:ISTQSLMTE 111
QY 132 ----NEPTESPANVDASVYFRKKWD:SKGDKKEYAELIRKQ 168
DB 112 NPYHNEPC-----FEQER-HPGDSKNYNECTRHE 139

RESULT 31
US-09-799-777-19
```

Sequence 19, Application US/9999777
Patent No. US20020091244A2

GENERAL INFORMATION:

APPLICANT: Ial, Preeti
Hillman, Jennifer L.
Corley, Neil C.
Gugler, Karl J.
Baugh, Mariah
Sathor, Susan
Shah, Puri

TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS

NUMBER OF SEQUENCES: 124

CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

APPLICATION DATA:

APPLICATION NUMBER: US/09/799,777
FILING DATE: 06-Mar-2001

CLASSIFICATION: <Unknown>

APPLICATION NUMBER: US/09/002,485

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: HILLINGS, LUCY J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0455 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 859-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 805 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: IHRNO103

CLONE: 1440381

SEQUENCE DESCRIPTION: SEQ ID NO: 19 :

US-09-749-777 19

Query Match: 11.6%; Score 175; DB 10; Length 805;
Best Local Similarity 25.0%; Pred. No. 3.4e-08;
Matches 79; Conservative 4; Mismatches 108; Indels 88; Gaps 14;

QY 1 MAQOMISSKALMLDKLSQEEPVERFTLVDSLDLYNWEVAIFGLPNTLYEGGYEKA 60

DE 526 LAGEAVILTSLSPLSSSSV-----FVRCDEERLQIMKVLITGVADTPYANGCFE 576

QY 61 HKEKIDYPSPTFRELKKMH-----PNLYENGQVCLSLRHPPVDDPQSGELPSEWN 115

DB 577 DVPYFPQYSPSPPLVNETTGHVSREPNLYNDGKVCLSL-----NTHGR-PEEKWN 630

QY 116 P-TONVRIILSV:SLK-----NPF-----NIFSPANYDASVMEFRKWRDSKGDKEYAE 163

DB 631 QTSSTPLQVIVSQSLILVAEPYFNPQYFHSKCI:PSCTQSS-----REYDG 677

QY 164 ITRKQVNTAEAKQGVKVPITLAEYCIKTKVPSNDNSDLYDLDYDDIDDEDEEE 223

DB 678 NTRQ --ATVKWMLQIRNPSV-----CFK-----EVIHKHFKYLKRVETMAQCEE 720

QY 224 ---DAKCYDDDDSCMS-----RDVLLQCPCTA-----LPSCAKRG 257

DB 721 WTAD:QYSSKRVKRGVGRMTSHHAAALKRHTAQLREELIKLQKPEGLQPTDDAPEVCRAIT 760

QY 258 ASGDAMAPQPKPIHS 273

DB 781 GABEILMHQVXPSSS 796

RESULT 32

US-09-925-302-746

Sequence 746, Application US/09925302

Patent No. US20020044941A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA104

CURRENT APPLICATION NUMBER: US/09/925,302

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05918

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 896

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 746

LENGTH: 186

TYPE: PRT

ORGANISM: Homo sapiens

US-09-925-302-746

Query Match: 11.2%; Score 169; DB 10; Length 186;

Best Local Similarity 31.0%; Pred. No. 1.7e-08;

Matches 40; Conservative 23; Mismatches 50; Indels 16; Gaps 4;

QY 38 LYNNFVAIFGLPNTLYEGGYEKAHKEFTDYPSPTFRLTKMHHPNIVE-NGDVCIS1 96

DB 52 LNEFVVKYGPQGIYEGGVKVKVDLPKAYPKSPSIGFMKKIFHPNIDEASIVCLDV 111

QY 97 LHPVDDPQSGELPSEWNPTQNVRTILLSVI-SLLNEPNTSPANDVASVME-RKWRDS 154

DB 112 ?-----NOTWALYDITNIFESFLQQLAYPNP:IDPLNGIAAAVYLRHPEY 158

QY 155 KGDKUEYAE 163

DB 159 KQTKKEYIQ 167

RESULT 33

US-10-157-669-1

Sequence 1, Application US/10157669

Publication No. US20030054385A1

GENERAL INFORMATION:

APPLICANT: Ial, Preeti G.

APPLICANT: Jackson, Jennifer L.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN UBIQUITIN-CONJUGATING ENZYMES

FILE REFERENCE: PF-0356-2 CIP

CURRENT APPLICATION NUMBER: US/10/157,669

CURRENT FILING DATE: 2002-05-28

PRIOR APPLICATION NUMBER: 09/520,076; 09/359,967; 08/965,689; 08/433,750

PRIOR FILING DATE: 2000-03-07; 1999-07-22; 1997-11-06; 1997-09-23

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PERL Program

SEQ ID NO 1

LENGTH: 185

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. US20030054385A1 001762CB1

US-10-157-669-1

Query Match

Best Local Similarity 10.7%; Score 161; DB 9; Length 185;

Matches 41; Conservative 39; Mismatches 69; Indels 12; Gaps 5;

QY 4 QOMTSOKALMLEKSLQDEPVEGRITLVDSGLYNNEVAIFGLPNT-LYEGGYEKAHI 62
 DB 28 RRVSVDRKLLVKEVAELANLPTCKVHFFDPNKLHCQJLV--TPDEGYOQCKQFET 85
 QY 63 KFPIDYPSPTFRFLTKMHPNINYEYDVCISILHPPVDDPQSGELPSERNWPTQVRT 122
 DB 86 EVDPAVMYPPVKVCKLTKIHPNITETGEICLSLL-----REHSDGTGMAPTRTKLD 138
 QY 123 ILLSVISLLNEPNTF-SPANVDASV-MFKKWRDSKGDKEY 161
 DB 139 VVWGLNSLFDLLNEDDPLNLEAAEHHLRDKEDFRNKVDY 179

RESULT 34

US-09-840-787-2
 : Sequence 2, Application US/09840787
 : Patent No. US20020058264A1
 : GENERAL INFORMATION:
 : APPLICANT: Lal, Freeti
 : Hillman, Jennifer L.
 : Bandman, Olga
 : Shah, Purvi
 : Au-Young, Janice
 : Yue, Henry
 : Guegler, Karl J.
 : Corley, Neil C.
 : TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
 : NUMBER OF SEQUENCES: 98
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Drive
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSEQ for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/840,787
 : FILING DATE: 23-Apr-2002
 : CLASSIFICATION: <Unknown>
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 09/518,865
 : FILING DATE: <Unknown>
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Billings, Lucy J.
 : REGISTRATION NUMBER: 36,745
 : REFERENCE/DOCKET NUMBER: PF-0356 US
 : TELEPHONE: 415-855-9555
 : TELEFAX: 415-845 4166
 : TELEX: <Unknown>
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 185 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : LIBRARY: U937N0101
 : CLONE: 1762
 : SEQUENCE DESCRIPTION: SEQ ID NO: 2 :

Query Match 10.7%, Score 161, DB 10; Length 185;
 Best Local Similarity 25.5%, Pred. No. 9.7e-08;
 Matches 41; Conservative 39; Mismatches 69; Indels 12; Gaps 5;

QY 4 QOMTSOKALMLEKSLQDEPVEGRITLVDSGLYNNEVAIFGLPNT-LYEGGYEKAHI 62

DB 28 RRVSVDRKLLVKEVAELANLPTCKVHFFDPNKLHCQJLV--TPDEGYOQCKQFET 85
 QY 63 KFPIDYPSPTFRFLTKMHPNINYEYDVCISILHPPVDDPQSGELPSERNWPTQVRT 122
 DB 86 EVDPAVMYPPVKVCKLTKIHPNITETGEICLSLL-----REHSDGTGMAPTRTKLD 138
 QY 123 ILLSVISLLNEPNTF-SPANVDASV-MFKKWRDSKGDKEY 161
 DB 139 VVWGLNSLFDLLNEDDPLNLEAAEHHLRDKEDFRNKVDY 179

RESULT 35

US-10-102-806-723
 : Sequence 723, Application US/10102806
 : Publication No. US20030054421A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 : FILE REFERENCE: PA103P1C1
 : CURRENT APPLICATION NUMBER: US/10/102,806
 : PRIOR FILING DATE: 2002-03-22
 : PRIOR APPLICATION NUMBER: 09/925,298
 : PRIOR FILING DATE: 2001-08-10
 : PRIOR APPLICATION NUMBER: PCT/US00/05881
 : PRIOR FILING DATE: 2000-03-08
 : PRIOR APPLICATION NUMBER: 60/124,270
 : PRIOR FILING DATE: 1999-03-12
 : NUMBER OF SEQ ID NOS: 846
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 723
 : LENGTH: 190
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-10-102-806-723

Query Match 10.6%, Score 159.5; DB 9; Length 190;
 Best Local Similarity 26.3%, Pred. No. 1.4e-07;
 Matches 51; Conservative 33; Mismatches 69; Indels 41; Gaps 7;

QY 8 SSOKALMLEKSLQDEPVE-----GFRITLVDSGLYNNEVAIFGLPNT-LYEGGYEKAH 61
 DB 30 SSKASAAQLR-IQKDIHNLNLPKTCDSFDDPLNFKLVI--CPDEGYFKSGKEVS 86
 QY 62 IKFPIDYPSPTFRFLTKMHPNINYEYDVCISILHPPVDDPQSGELPSERNWPTQVNR 121
 DB 87 FKVGCGYPHDPKVKCTMYHPNIDLEGWVCLNIIR-----EDMKPVLITIN 133
 QY 122 TILSVISLLNEPNTFSPANVDASVMFKKWRDSKGDKEYAEIIR--KQVSATKAEAKD 179
 DB 134 SILVGLQYLFLPNPDPDLN-----KEAFVLQNNRRFLPQNVQSRMK 176
 QY 180 GVKVPTTLAAYCIK 193
 DB 177 GGYIGSYFERCLK 190

RESULT 36

US-09-925-302-620
 : Sequence 620, Application US/09925302
 : Patent No. US20020044941A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 : FILE REFERENCE: PA104
 : CURRENT APPLICATION NUMBER: US/09/925,302
 : CURRENT FILING DATE: 2001-08-10
 : PRIOR APPLICATION NUMBER: PCT/US00/05918
 : PRIOR FILING DATE: 2000-03-08
 : PRIOR APPLICATION NUMBER: 60/124,270
 : PRIOR FILING DATE: 1999-03-12
 : NUMBER OF SEQ ID NOS: 896
 : SOFTWARE: PatentIn Ver. 2.0


```

? SEQ ID NO 620
? LENGTH: 305
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (20)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (216)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-99-925-302-620

Query Match 10.58; Score 158; DB 10; Length 305;
Best Local Similarity 25.14; Pred. No. 3.7e-07;
Matches 66; Conservative 34; Mismatches 93; Indels 70; Gaps 12:

Qy 35 ESLEYNKEVAHGLNLTLYEGYEKAKIKFPIDYGYSTPIERFLTAKWHHNIYENG-DYC 93
Db 22 EDNIFWHTVKGPPDSDIFGAGVTHGR:VLPPEYPMKPPS:ILITANGR---FEVKKKIC 78
Qy 94 ISILHPFVDEPQSGELPSERNPNTQNVRTILLSVI-----SLL 131
Db 79 ISI--- ---SCHHP-EWQSWSIKRTALLAI:CFWPKGRGAIGSLDY:TFERRALA 128
Qy 132 NEHNP-----SPANVDASVWFKNWFSKQKQKEYAEI:RKQVSAI:KAAEAKDGVKVPPT 186
Db 129 KKSQDCCCECGGSAMKQVLLPLKSGSDSQADQEAKE:EAR-QIS-EKAEVNSRG----- 180
Qy 187 LAECG:IKWVQSNENSSDLYLDLYDDDDDEHFEADYDQDDDSUMSRDVLILQCP 246
Db 181 -----KTSISDLNHSPT:THD---LQDDIPTI:QGATAS:SYGXNSNSAAS----- 223
Qy 247 TALPSQAKGKSGSLAMAPQOK 269
Db 224 FHQPIQPVAKNTS-----MSPKQR 242

RESULT 37
us-09-785-671-7
? Sequence 7, Application us/09785671
? Patent No. US20020128169A1
? GENERAL INFORMATION:
? APPLICANT: Kuo, Choon-Joo
? APPLICANT: Lee, Mu-Ea
? APPLICANT: Haber, Edgar
? TITLE OF INVENTION: UNIFICATION OF THE TRANSCRIPTION FACTOR E2A
? FILE REFERENCE: 05434-035002
? CURRENT APPLICATION NUMBER: US/09/785,671
? CURRENT FILING DATE: 2001-09-14
? PRIOR APPLICATION NUMBER: US 08/986,605
? PRIOR FILING DATE: 1997-12-08
? PRIOR APPLICATION NUMBER: US 08/825,476
? PRIOR FILING DATE: 1997-03-28
? PRIOR APPLICATION NUMBER: US 60/014,388
? PRIOR FILING DATE: 1996-03-28
? NUMBER OF SEQ ID NOS: 7
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 7
? LENGTH: 157
? TYPE: PRT
? ORGANISM: Artificial sequence
? FEATURE:
? OTHER INFORMATION: Consensus sequence
? NAME/KEY: VARIANT
? LOCATION: (3)...(3)
? OTHER INFORMATION: Xaa - Ser or Gly
? NAME/KEY: VARIANT
? LOCATION: (4)...(4)
? OTHER INFORMATION: Xaa - Leu or Ile
? NAME/KEY: VARIANT
? LOCATION: (5)...(5)
? OTHER INFORMATION: Xaa - Cys or Ala
? NAME/KEY: VARIANT
? LOCATION: (7)...(7)
? OTHER INFORMATION: Xaa - Gln or Ser
? NAME/KEY: VARIANT
? LOCATION: (10)...(10)
? OTHER INFORMATION: Xaa - Gln or Ala
? NAME/KEY: VARIANT
? LOCATION: (11)...(11)
? OTHER INFORMATION: Xaa - Glu or Gln
? NAME/KEY: VARIANT
? LOCATION: (15)...(15)
? OTHER INFORMATION: Xaa - Lys or Ala
? NAME/KEY: VARIANT
? LOCATION: (25)...(25)
? OTHER INFORMATION: Xaa - Tyr or Val
? NAME/KEY: VARIANT
? LOCATION: (27)...(27)
? OTHER INFORMATION: Xaa - Lys or Val
? NAME/KEY: VARIANT
? LOCATION: (29)...(29)
? OTHER INFORMATION: Xaa - Val or Thr
? NAME/KEY: VARIANT
? LOCATION: (31)...(31)
? OTHER INFORMATION: Xaa - Lys or Asn
? NAME/KEY: VARIANT
? LOCATION: (32)...(32)
? OTHER INFORMATION: Xaa - Ala or Pro
? NAME/KEY: VARIANT
? LOCATION: (35)...(35)
? OTHER INFORMATION: Xaa - Ser or Thr
? NAME/KEY: VARIANT
? LOCATION: (37)...(37)
? OTHER INFORMATION: Xaa - Asp or Asn
? NAME/KEY: VARIANT
? LOCATION: (39)...(39)
? OTHER INFORMATION: Xaa - Gln or Met
? NAME/KEY: VARIANT
? LOCATION: (40)...(40)
? OTHER INFORMATION: Xaa - Lys or Asn
? NAME/KEY: VARIANT
? LOCATION: (43)...(43)
? OTHER INFORMATION: Xaa - Ala or Cys
? NAME/KEY: VARIANT
? LOCATION: (44)...(44)
? OTHER INFORMATION: Xaa - Gly or Ala
? NAME/KEY: VARIANT
? LOCATION: (49)...(49)
? OTHER INFORMATION: Xaa - Glu or Lys
? NAME/KEY: VARIANT
? LOCATION: (52)...(52)
? OTHER INFORMATION: Xaa - Asn or Pro
? NAME/KEY: VARIANT
? LOCATION: (54)...(54)
? OTHER INFORMATION: Xaa - Ala or Glu
? NAME/KEY: VARIANT
? LOCATION: (57)...(57)
? OTHER INFORMATION: Xaa - Val or Leu
? NAME/KEY: VARIANT
? LOCATION: (58)...(58)
? OTHER INFORMATION: Xaa - Tyr or Phe
? NAME/KEY: VARIANT
? LOCATION: (59)...(59)
? OTHER INFORMATION: Xaa - Pro or Lys
? NAME/KEY: VARIANT
? LOCATION: (60)...(60)
? OTHER INFORMATION: Xaa - Ile or Leu
? NAME/KEY: VARIANT
? LOCATION: (61)...(61)
? OTHER INFORMATION: Xaa - Thr or Arg
? NAME/KEY: VARIANT
? LOCATION: (62)...(62)
? OTHER INFORMATION: Xaa - Val or Met
? NAME/KEY: VARIANT

```

LOCATION: (63)...(63)
OTHER INFORMATION: Xaa - Glu or Leu
NAME/KEY: VARIANT
LOCATION: (64)...(64)
OTHER INFORMATION: Xaa - Tyr or Phe
NAME/KEY: VARIANT
LOCATION: (65)...(65)
OTHER INFORMATION: Xaa - Pro or Lys
NAME/KEY: VARIANT
LOCATION: (66)...(66)
OTHER INFORMATION: Xaa - Asn or Asp
NAME/KEY: VARIANT
LOCATION: (67)...(67)
OTHER INFORMATION: Xaa - Glu or Asp
NAME/KEY: VARIANT
LOCATION: (71)...(71)
OTHER INFORMATION: Xaa - Lys or Ser
NAME/KEY: VARIANT
LOCATION: (75)...(75)
OTHER INFORMATION: Xaa - Val or Cys
NAME/KEY: VARIANT
LOCATION: (78)...(78)
OTHER INFORMATION: Xaa - Pro or Glu
NAME/KEY: VARIANT
LOCATION: (79)...(79)
OTHER INFORMATION: Xaa - Ala or Pro
NAME/KEY: VARIANT
LOCATION: (80)...(80)
OTHER INFORMATION: Xaa - Gly or Pro
NAME/KEY: VARIANT
LOCATION: (81)...(81)
OTHER INFORMATION: Xaa - Phe or Leu
NAME/KEY: VARIANT
LOCATION: (82)...(82)
OTHER INFORMATION: Xaa - Tyr or Phe
NAME/KEY: VARIANT
LOCATION: (92)...(92)
OTHER INFORMATION: Xaa - Ile or Val
NAME/KEY: VARIANT
LOCATION: (98)...(98)
OTHER INFORMATION: Xaa - Asn or Glu
NAME/KEY: VARIANT
LOCATION: (101)...(101)
OTHER INFORMATION: Xaa - Gln or Lys
NAME/KEY: VARIANT
LOCATION: (109)...(109)
OTHER INFORMATION: Xaa - Leu or Ile
NAME/KEY: VARIANT
LOCATION: (113)...(113)
OTHER INFORMATION: Xaa - Val or Leu
NAME/KEY: VARIANT
LOCATION: (116)...(116)
OTHER INFORMATION: Xaa - Val or Ile
NAME/KEY: VARIANT
LOCATION: (118)...(118)
OTHER INFORMATION: Xaa - Asp or Glu
NAME/KEY: VARIANT
LOCATION: (121)...(121)
OTHER INFORMATION: Xaa - Asp or Asn
NAME/KEY: VARIANT
LOCATION: (122)...(122)
OTHER INFORMATION: Xaa - Ser or Glu
NAME/KEY: VARIANT
LOCATION: (125)...(125)
OTHER INFORMATION: Xaa - Pro or Ile
NAME/KEY: VARIANT
LOCATION: (126)...(126)
OTHER INFORMATION: Xaa - Asn or Gln
NAME/KEY: VARIANT
LOCATION: (127)...(127)
OTHER INFORMATION: Xaa - Ser or Asp
NAME/KEY: VARIANT
LOCATION: (131)...(131)

OTHER INFORMATION: Xaa - Glu or Ala
NAME/KEY: VARIANT
LOCATION: (132)...(132)
OTHER INFORMATION: Xaa - Pro or Glu
NAME/KEY: VARIANT
LOCATION: (134)...(134)
OTHER INFORMATION: Xaa - Trp or Tyr
NAME/KEY: VARIANT
LOCATION: (135)...(135)
OTHER INFORMATION: Xaa - Arg or Thr
NAME/KEY: VARIANT
LOCATION: (136)...(136)
OTHER INFORMATION: Xaa - Ser or Ile
NAME/KEY: VARIANT
LOCATION: (137)...(137)
OTHER INFORMATION: Xaa - Phe or Tyr
NAME/KEY: VARIANT
LOCATION: (138)...(138)
OTHER INFORMATION: Xaa - Ser or Cys
NAME/KEY: VARIANT
LOCATION: (139)...(139)

Query Match 9.88; Score 148.5; DB 10; Length 157;

Best Local Similarity 29.48; Pred. No. 1.2e-06;
Matches 40; Conservative 9; Mismatches 72; Indels 15; Gaps 2;

Qy 13 LMLEKSLQEEPEVGEFRILLVDESD----LYNWEVAIEGLPNILYEGGYEKAKHIFPIDY 68
Db 9 LXXERXXWKDHPEGFAXAXKXXDXMXLXXWEXXIIGKXGTWXXGXXXXXXXYY 58
Qy 69 PYSPTIEFLTKMHPNIYENGWVCISILHPPVDPQSGELPSERNWPFQNVRIE LLSV 128
Db 69 PSAPPKXKEXXXHPNVPSGTCLSLKEXD-----WEPATTKQIXLXQ 117
Qy 129 SILNEPTESPAVDA 144
Db 118 XLIXPNXXXXPAQXXA 133

RESULT 38

US-09-864-761-35682
Sequence 35682; Application US/0986476;
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCI/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCI/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCI/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCI/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCI/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/638,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49/117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 35682
 LENGTH: 55
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AP00553.1
 OTHER INFORMATION: EXPRESSED IN B1474, SIGNAL - 3.3
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 1.4
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 2
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 1.9
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 1.1
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1.4
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 1.6
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 7.8
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL - 5
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 0.99
 OTHER INFORMATION: EST_HUMAN HIL: HIL76937.1, EVALUATION 3.00e-16
 OTHER INFORMATION: SWISSPROT HIL: P51966, EVALUATION 6.00e-17
 US-09-864-761-35682

Query Match 8.9%; Score 134; DB 10; Length 55;
 Best Local Similarity 37.7%; Pred. No. 6.8e-06;
 Matches 23; Conservative 9; Mismatches 17; Indels 12; Gaps 1;

QY 56 GYFKAHKEPDIYSPPTFRILKMHPIYENGDCVCSILHPVDDPQSGELSPERN 115
 DB 1 GAFREINFPAYRKPPTKTEKTIYHNIDKQVCLIV-----ISAENWK 48
 QY 116 P 116
 DB 49 P 49

RESULT 35
 US-09-884-720-16
 Sequence 16, Application US/09883720
 Patent No. US2002022256A1
 GENERAL INFORMATION:
 APPLICANT: Haldwin, Donald A.
 APPLICANT: Briggs, Steven P.
 APPLICANT: Crane, Virginia C.
 TITLE OF INVENTION: Maltose Histone Deacetylases And Their Uses
 FILE REFERENCE: 5718 44,
 CURRENT APPLICATION NUMBER: US/09/883,720
 PRIOR FILING DATE: 2001-06-18
 PRIOR APPLICATION NUMBER: 09/282,305
 PRIOR FILING DATE: 1999-03-31
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 16
 LENGTH: 311
 TYPE: PRT
 ORGANISM: Hec mays
 US-09-883-720-16

Query Match 6.9%; Score 104.5; DB 10; Length 311;
 Best Local Similarity 25.1%; Pred. No. 0.046;
 Matches 44; Conservative 25; Mismatches 68; Indels 37; Gaps 6;
 QY 130 LLINEPTESPANVDASVMPKWR-----DSKGDKEY-AEIIKOVSTAKAA 176
 DB 69 VENKEFELSHTSKTSIVFFSGYKVEQPIEGDEMOLDSEDEEELNIPVIKENGAKDKEE 128
 QY 177 EKGVK-VPTTIAEYGLTKVPSNDSLLYDDIYDDIDUDEDEEE-----DAD 226
 DB 129 QKNOEKAATAASKSSLCLEKSKDDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSD 189
 QY 227 CYDDDDSDMSRSDVLLQPCPTALPSQAKSGASGDL-----AMAPQOK 269
 DB 189 SSEDUT---SDDDEETPTPKKPEAGKKGKAENALKTPILSDKKKAKVATPPACK 239

RESULT 40
 US-09-925-302-537
 Sequence 537, Application US/09925302
 Patent No. US20020044941A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 FILE REFERENCE: PA104
 CURRENT APPLICATION NUMBER: US/09/925,302
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05918
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 896
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 537
 LENGTH: 241
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-925-302-537

Query Match 6.8%; Score 103; DB 10; Length 241;
 Best Local Similarity 26.7%; Pred. No. 0.046;
 Matches 40; Conservative 25; Mismatches 51; Indels 34; Gaps 8;
 QY 12 ALMLEKLSQEPVEGERITLVDSDLYNWEVAIFGLPNTLYEGYFKAHKFPIDYPS 71
 DB 26 SLAAETIVVVKLPQVYVQPSVSALM-WGVIF-IRHCLYQDGVFKETVYIPUNYD-- 81
 QY 72 PPTFRILKMHPIYENGDCV-----ISILHPVDDPQSGELSPER-----WNPTQN-V 120
 DB 82 -----DGD-CRLVFDIPVHPLV-DPTSGELDKRAFAKWRNRNHHI 122
 QY 121 RTILLSVTSILNEPTESPANVDASVMPK 150
 DB 123 WOVLVYARVFKYKIDTASPLNPEANVLYEK 152

Search completed: April 10, 2003, 10:37:44
 Job time: 16.6118 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:29:45 ; Search time 15.2612 Seconds
(without alignments)
543.684 Million cell updates

Title: US-09-930-026-2

Perfect score:

Sequence: 1 MAQQQMTSSQKALMLELKS.....AMAPQOKPIISGWNTHSSC 282

Scoring table: BI.OSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 135 summaries

Database : Issued Patents A: *

1: /cqn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cqn2_6/plodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pcp:★

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.per:*

6: /cqn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	Time				
1	1511	100.0	282	4	US-09-058	368-2	Sequence 2, Appli
2	1010	66.8	298	3	US-08-767	-942A-25	Sequence 25, Appl
3	1010	66.8	298	4	US-09-177	-165A-23	Sequence 23, Appl
4	460.5	30.5	170	2	US-08-820	-170A-22	Sequence 22, Appl
5	460.5	30.5	170	2	US-08-918	-723-1	Sequence 1, Appl
6	460.5	30.5	170	2	US-09-237	-507-1	Sequence 1, Appl
7	460.5	30.5	170	3	US-09-055	-699-22	Sequence 22, Appl
8	460.5	30.5	170	4	US-09-273	-565-22	Sequence 22, Appl
9	460.5	30.5	170	4	US-09-565	-538-22	Sequence 22, Appl
10	460.5	30.5	170	4	US-09-561	-468-22	Sequence 22, Appl
11	450	29.8	166	2	US-08-918	-723-4	Sequence 4, Appl
12	450	29.8	166	2	US-09-237	-507-4	Sequence 4, Appl
13	432.5	28.6	295	2	US-08-679	-765-5	Sequence 5, Appl
14	432.5	28.6	295	2	US-09-106	-525-5	Sequence 5, Appl
15	432.5	28.6	295	4	US-09-318	-317-5	Sequence 5, Appl
16	432.5	28.6	295	4	US-09-177	-165A-22	Sequence 22, Appl
17	428.5	28.4	164	2	US-08-918	-723-3	Sequence 3, Appl
18	428.5	28.4	164	2	US-09-237	-507-3	Sequence 3, Appl
19	380.5	25.2	165	3	US-08-965	-689A-3	Sequence 3, Appl
20	380.5	25.2	165	4	US-09-339	-967-3	Sequence 3, Appl
21	358.5	23.7	165	3	US-08-965	-689A-6	Sequence 6, Appl
22	358.5	23.7	165	4	US-09-359	-967-6	Sequence 6, Appl
23	310	20.5	172	3	US-09-358	-580-14	Sequence 14, Appl
24	308	20.4	172	1	US-08-318	-947A-11	Sequence 11, Appl
25	308	20.4	172	2	US-08-795	-303-11	Sequence 11, Appl
26	295	19.5	152	2	US-08-247	-904B-16	Sequence 16, Appl
27	295	19.5	152	3	US-08-767	-942A-17	Sequence 17, Appl


```

: RESULT 5
: US-08-918-723-1
: : Sequence 1, Application US/08918723
: : Patent No. 5863779
: : GENERAL INFORMATION:
: : APPLICANT: lai, preeti
: : APPLICANT: Corley, Neil C.
: : TITLE OF INVENTION: UBC7-LIKE URICOTIN-CONJUGATING ENZYME
: : NUMBER OF SEQUENCES: 4
: : CORRESPONDENCE ADDRESS:

```


APPLICATION NUMBER: 08/820,170
FILING DATE:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-055-639-22

Query Match 30.5%: Score 460.5; DB 3; Length 170;
Best Local Similarity 52.4%: Pred. No. 3.4e-40;
Matches 87; Conservative 28; Mismatches 44; Indels 7; Gaps 3;
QY 6 MTSSOKALML--ELKSLQEPVGGFRITLVDSLYNNEVAIFGLPNTLYEGGYKKAHLK 63
DB 1 MTEQSALLRRLQAEIKNKPVGFSAGLIIDNDLYRWEVLIIIGPDDTLYEGGYKKAHLT 60
QY 64 FPIIDYPSPTFRFLTKMHPNLYENGVCISILHPVDVDDQSGEILPSPERNPTQNVRTI 123
DB 61 FPKDYLPRPPKMKVITEIHPNVDRKNGDVCISILHPGDKYGYKPEERWLPITHIVETI 120
QY 124 LLSVISLLENPFTSPANDASVMFKWRDSKGD--KEYAEIIRK 167
DB 121 MISVLSMLADPNQDSANVDAA---KWRDRNGEKKKVARCVRK 163

RESULT 8

US-09-273-565-22
Sequence 22, Application US/09273565A
Patent No. 6166190
GENERAL INFORMATION:
APPLICANT: FUJIWARA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
FILE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/273,565A
EARLIER FILING DATE: 1999-03-22
EARLIER APPLICATION NUMBER: 09/055,699
EARLIER FILING DATE: 1998-04-07
EARLIER APPLICATION NUMBER: 08/820,170
EARLIER FILING DATE: 1997-03-19
EARLIER APPLICATION NUMBER: JP 63410/1996
EARLIER FILING DATE: 1996-03-19
EARLIER APPLICATION NUMBER: JP 69163/1997
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 22
LENGTH: 170
TYPE: PRT
ORGANISM: Homo sapiens
US-09-273-565-22

Query Match 30.5%: Score 460.5; DB 4; Length 170;
Best Local Similarity 52.4%: Pred. No. 3.4e-40;
Matches 87; Conservative 28; Mismatches 44; Indels 7; Gaps 3;
QY 6 MTSSOKALML--ELKSLQEPVGGFRITLVDSLYNNEVAIFGLPNTLYEGGYKKAHLK 63
DB 1 MTEQSALLRRLQAEIKNKPVGFSAGLIIDNDLYRWEVLIIIGPDDTLYEGGYKKAHLT 60
QY 64 FPIIDYPSPTFRFLTKMHPNLYENGVCISILHPVDVDDQSGEILPSPERNPTQNVRTI 123
DB 61 FPKDYLPRPPKMKVITEIHPNVDRKNGDVCISILHPGDKYGYKPEERWLPITHIVETI 120
QY 124 LLSVISLLENPFTSPANDASVMFKWRDSKGD--KEYAEIIRK 167

Db 121 MISVLSMLADPNQDSANVDAA---KWRDRNGEKKKVARCVRK 163

RESULT 9

US-09-565-538-22
Sequence 22, Application US/09565538
Patent No. 6333404
GENERAL INFORMATION:
APPLICANT: FUJIWARA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
FILE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/565,538
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 09/273,565
PRIOR FILING DATE: 1999-03-22
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 69163/1997
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 22
LENGTH: 170
TYPE: PRT
ORGANISM: Homo sapiens
US-09-565-538-22

Query Match 30.5%: Score 460.5; DB 4; Length 170;
Best Local Similarity 52.4%: Pred. No. 3.4e-40;
Matches 87; Conservative 28; Mismatches 44; Indels 7; Gaps 3;
QY 6 MTSSOKALML--ELKSLQEPVGGFRITLVDSLYNNEVAIFGLPNTLYEGGYKKAHLK 63
DB 1 MTEQSALLRRLQAEIKNKPVGFSAGLIIDNDLYRWEVLIIIGPDDTLYEGGYKKAHLT 60
QY 64 FPIIDYPSPTFRFLTKMHPNLYENGVCISILHPVDVDDQSGEILPSPERNPTQNVRTI 123
DB 61 FPKDYLPRPPKMKVITEIHPNVDRKNGDVCISILHPGDKYGYKPEERWLPITHIVETI 120
QY 124 LLSVISLLENPFTSPANDASVMFKWRDSKGD--KEYAEIIRK 167
DB 121 MISVLSMLADPNQDSANVDAA---KWRDRNGEKKKVARCVRK 163

RESULT 10

US-09-661-468-22
Sequence 22, Application US/09661468
Patent No. 6376189
GENERAL INFORMATION:
APPLICANT: FUJIWARA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
FILE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/661,468
CURRENT FILING DATE: 2000-09-13
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 69163/1997
PRIOR FILING DATE: 1997-03-05

Db 62 QNYPSPPTVRETSMHNPVNSGRVCISILHPPCDPSGYELASERTPWIIHVESIML 121
 QY 126 SVISLINPENTSPANDASVMPKWKWDKSKG-KREYAEH1RK 167
 Db 122 SLISMLSGNDESPANFAA---KEWRDKRDEKFKKVSRCVRK 161

RESULT 13

US-08-679-765-5
 ; Sequence 5, Application US/08679765
 ; Patent No. 5840866
 ; GENERAL INFORMATION:
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Goli, Surya K.
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: A NOVEL HUMAN UBIQUITIN-CONJUGATING ENZYME
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: U.S.
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/679,765
 ; FILING DATE: Filed Herewith
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0093 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 295 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 788905
 ; US-08 679-765-5

Query Match 28.68; Score 432.5; DB 2; Length 295;
 Best Local Similarity 33.38; Pred. NC. 6.4e-37;
 Matches 93; Conservative 61; Mismatches 78; Indels 47; Gaps 81

QY 1 MAQOQMTSSOKALMLELKSLOE--EPVEGRITLVDSGLYNWEVAITSLPMTLYEGGY 57
 Db 1 MSSRKSTAS-SULLRQYRELIDPKKAIPSPHIEDDSNIFTWNIQVMVINEHSIYHGGF 59

QY 58 FKAHIFPTIDYPSPTTFRFTKMMHNTYENGDCVTSILHPPVDVDSGSGELPSERNPT 117
 Db 60 FKAQMRFPEDFPSPQFRFTPAIYHPNVYRGRICISILHQS-GDPMTEPDAETWSPV 118

QY 118 QNVRILLSVISLILNEPTFSPANVDASVMPKWKWDKSKGKAEIIRKONSATKAE 177
 Db 119 QTVESVLISVLSLEDPNINSANVDAVDYRK-----NPEQYKRVKMEVRSKODIP 172

QY 178 KGVKVPPTILAEYCITK-----VPSNDSGLLYDOLYDD----- 213
 Db 173 K-GFIMPTFSAYISQSKLDPEPSNKMDAUNFWYDSLDLDDDDGSGVILQDDYDGNHHI 231

QY 214 -----DIDDEDEEEEDACDYDDDSGMRSDVL 241

Db 232 PFEDZDVVYNDNDUUEETFEFDGDDDDDDNDSDNSVM 270

RESULT 14

US-09-196-525-5
 ; Sequence 5, Application US/09196525
 ; Patent No. 5989883
 ; GENERAL INFORMATION:
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Goli, Surya K.
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: A NOVEL HUMAN UBIQUITIN-CONJUGATING ENZYME
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: U.S.
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/196,525
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/679,765
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0093 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 295 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 788905
 ; US-09-196-525-5

Query Match 28.68; Score 432.5; DB 2; Length 295;
 Best Local Similarity 33.38; Pred. NC. 6.4e-37;
 Matches 93; Conservative 61; Mismatches 78; Indels 47; Gaps 81

QY 1 MAQOQMTSSOKALMLELKSLOE--EPVEGRITLVDSGLYNWEVAITSLPMTLYEGGY 57
 Db 1 MSSRKSTAS-SULLRQYRELIDPKKAIPSPHIEDDSNIFTWNIQVMVINEHSIYHGGF 59

QY 58 FKAHIFPTIDYPSPTTFRFTKMMHNTYENGDCVTSILHPPVDVDSGSGELPSERNPT 117
 Db 60 FKAQMRFPEDFPSPQFRFTPAIYHPNVYRGRICISILHQS-GDPMTEPDAETWSPV 118

QY 118 QNVRILLSVISLILNEPTFSPANVDASVMPKWKWDKSKGKAEIIRKONSATKAE 177
 Db 119 QTVESVLISVLSLEDPNINSANVDAVDYRK-----NPEQYKRVKMEVRSKODIP 172

QY 178 KGVKVPPTILAEYCITK-----VPSNDSGLLYDOLYDD----- 213
 Db 173 K-GFIMPTFSAYISQSKLDPEPSNKMDAUNFWYDSLDLDDDDGSGVILQDDYDGNHHI 231

QY 214 -----DIDDEDEEEEDACDYDDDSGMRSDVL 241

Db 232 PFEDDDVYNNDDDDRIEFEDDDDDDDSDNDSDVM 270

RESULT 15
US-09-318-317-5
Sequence 5, Application US/09318317
Patent No. 6172199
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: A NOVEL HUMAN UBIQUITIN-CONJUGATING
TITLE OF INVENTION: ENZYME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/318,317
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/196,525
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0093 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 295 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 788905
US-09-318-317-5

Query Match 28.6%, Score 432.5; DB 4; Length 295;

Best Local Similarity 33.3%; Pred. No. 6.4e-37;

Matches 93; Conservative 61; Mismatches 78; Indels 47; Gaps 8;

QY 1 MAQQMTSSKALMLEKSLQ--EPVEGRITLVDSLYNWEVAIFGL-PNTLYPGGY 57

Db 1 MSSRKSTAS-SLLRGYRELTPKKAIPSEHIELEDDSNFTWNGVWLNEDSIYHG 59

QY 58 FKAHKEPTDYPSPPTFRFTKMMHNPVYNGVCTSLHPTVDPOSGLPSEKWNPT 117

Db 60 FKAQRPEDPFPSPQFPPTAIYHPNVYDGRICISLHQS GDMPTDFDAFTWSPV 118

QY 118 QNVRTLLSVLSLNEPNTFSPANVDASVMEKRWKSGKDEYAEIIRKOVSAKAEAE 177

Db 119 QTVESVLISVLSLEDPNLSNPANDVAADYRK-----NPFQYKQKMEVRSKODIP 172

QY 178 KGVKVPPTLAECYKIK-----VPSNDSSDLYDDYD----- 213

Db 173 K-GFIMPTSESAYISUSKLDPESKMDADNFWYDSLDGDSGVILGDDYDGNHHI 231

QY 214 -----DIDUDDFEEDADCYDDDDSGHRSRDL 241

Db 232 PFEDDDVYNNDDDDRIEFEDDDDDDDSDNDSDVM 270

Db 232 PFEDDDVYNNDDDDRIEFEDDDDDDDSDNDSDVM 270

RESULT 16
US-09-177-165A-22
Sequence 22, Application US/09177165A
Patent No. 6426205
GENERAL INFORMATION:
APPLICANT: Tyers, Mike
APPLICANT: Willem, Andrew
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
TITLE OF INVENTION: DEPENDENT PROTEOLYSIS
FILE REFERENCE: 11757.10US01
CURRENT APPLICATION NUMBER: US/09/177,165A
CURRENT FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/042,443
PRIOR FILING DATE: 1998-07-10
PRIOR APPLICATION NUMBER: 60/063,254
PRIOR FILING DATE: 1997-10-24
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 22
LENGTH: 295
TYPE: PRI
ORGANISM: Saccharomyces cerevisiae
US-09-177-165A-22

Query Match 26.6%, Score 432.5; DB 4; Length 295;

Best Local Similarity 33.3%; Pred. No. 6.4e-37;

Matches 93; Conservative 61; Mismatches 78; Indels 47; Gaps 8;

QY 1 MAQQMTSSKALMLEKSLQ--EPVEGRITLVDSLYNWEVAIFGL-PNTLYPGGY 57

Db 1 MSSRKSTAS-SLLRGYRELTPKKAIPSEHIELEDDSNFTWNGVWLNEDSIYHG 59

QY 58 FKAHKEPTDYPSPPTFRFTKMMHNPVYNGVCTSLHPTVDPOSGLPSEKWNPT 117

Db 60 FKAQRPEDPFPSPQFPPTAIYHPNVYDGRICISLHQS GDMPTDFDAFTWSPV 118

QY 118 QNVRTLLSVLSLNEPNTFSPANVDASVMEKRWKSGKDEYAEIIRKOVSAKAEAE 177

Db 119 QTVESVLISVLSLEDPNLSNPANDVAADYRK-----NPFQYKQKMEVRSKODIP 172

QY 178 KGVKVPPTLAECYKIK-----VPSNDSSDLYDDYD----- 213

Db 173 K-GFIMPTSESAYISUSKLDPESKMDADNFWYDSLDGDSGVILGDDYDGNHHI 231

QY 214 -----DIDUDDFEEDADCYDDDDSGHRSRDL 241

Db 232 PFEDDDVYNNDDDDRIEFEDDDDDDDSDNDSDVM 270

Query Match 28.6%, Score 432.5; DB 4; Length 295;

Best Local Similarity 33.3%; Pred. No. 6.4e-37;

Matches 93; Conservative 61; Mismatches 78; Indels 47; Gaps 8;

QY 1 MAQQMTSSKALMLEKSLQ--EPVEGRITLVDSLYNWEVAIFGL-PNTLYPGGY 57

Db 1 MSSRKSTAS-SLLRGYRELTPKKAIPSEHIELEDDSNFTWNGVWLNEDSIYHG 59

QY 58 FKAHKEPTDYPSPPTFRFTKMMHNPVYNGVCTSLHPTVDPOSGLPSEKWNPT 117

Db 60 FKAQRPEDPFPSPQFPPTAIYHPNVYDGRICISLHQS GDMPTDFDAFTWSPV 118

QY 118 QNVRTLLSVLSLNEPNTFSPANVDASVMEKRWKSGKDEYAEIIRKOVSAKAEAE 177

Db 119 QTVESVLISVLSLEDPNLSNPANDVAADYRK-----NPFQYKQKMEVRSKODIP 172

QY 178 KGVKVPPTLAECYKIK-----VPSNDSSDLYDDYD----- 213

Db 173 K-GFIMPTSESAYISUSKLDPESKMDADNFWYDSLDGDSGVILGDDYDGNHHI 231

QY 214 -----DIDUDDFEEDADCYDDDDSGHRSRDL 241

Db 232 PFEDDDVYNNDDDDRIEFEDDDDDDDSDNDSDVM 270

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,723
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0368 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 164 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 297965
US-08-918-723-3

Query Match 28.4% Score 428.5; DB 2; Length 164;

Best Local Similarity 48.0% Pred. No. 6,9e-37;

Matches 83; Conservative 25; Mismatches 56; Indels 9; Gaps 1;

QY 6 MTSOKALMELKSLQEPVGFRTILVDESDIYNWEVAIFGLPNTLYHGGYFKAHFKFP 65
DB 1 MEQSULLKKQLADMRKVPVDSAGLVDDNDIYKWEVVLGPPDLYEGGFEKALIDFP 60
QY 66 IDYPSPPTFRFLTKMHPNIYENGVCISILHPPVDDPSGELPSEKRNPNQVTRTILL 125
DB 61 RYQPKPKMKFISEIWHPNIDKGNVCISILHPPGDDKNGYERPFERFLPVHTVTETILL 120
QY 126 SVISLLNPNFTSPANDASVMEKRWDSKCKDEYAEIIRKQVSATKAEEK 178
DB 121 SVISMLTDPNFSPANVDAKMQR-----ENYAEKKKVAQCVRSQEE 164

RESULT 18

US-09-237-507-3

Sequence 3, Application US/09237507

Patent No. 5952181

GENERAL INFORMATION:

APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: UBC7-LIKE UBIQUITIN-CONJUGATING ENZYME
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,507
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/918,723

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0368 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 164 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 297965
US-09-237-507-3

Query Match 28.4% Score 428.5; DB 2; Length 164;

Best Local Similarity 48.0% Pred. No. 6,9e-37;

Matches 83; Conservative 25; Mismatches 56; Indels 9; Gaps 1;

QY 6 MTSOKALMELKSLQEPVGFRTILVDESDIYNWEVAIFGLPNTLYHGGYFKAHFKFP 65
DB 1 MEQSULLKKQLADMRKVPVDSAGLVDDNDIYKWEVVLGPPDLYEGGFEKALIDFP 60
QY 66 IDYPSPPTFRFLTKMHPNIYENGVCISILHPPVDDPSGELPSEKRNPNQVTRTILL 125
DB 61 RYQPKPKMKFISEIWHPNIDKGNVCISILHPPGDDKNGYERPFERFLPVHTVTETILL 120
QY 126 SVISLLNPNFTSPANDASVMEKRWDSKCKDEYAEIIRKQVSATKAEEK 178
DB 121 SVISMLTDPNFSPANVDAKMQR-----ENYAEKKKVAQCVRSQEE 164

RESULT 19

US-09-965-689A-3

Sequence 3, Application US/08965689A

Patent No. 6015702

GENERAL INFORMATION:

APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN UBIQUITIN CONJUGATING ENZYMS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,689A
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

```

: LENGTH: 165 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: ENDANOT01
: CLONE: 2456290

```

```

Query Match      25.24; Score 380.5; DB 3; Length 165;
Best local Similarity 50.34; Proc. NO. 6.9e-32;
Matches 80; Conservative 18; Mismatches 54; Indels 7; Gaps 2;

QY      11 KALMLELKSLOEPEVEGRITLVDESILYNKEVAIEGLPNTLYEGGYEKAIHKEPDIYYP 70
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       7 KRLMAEYKQILTPPGEGIVAGMMNEENFFWEALIMGPEITCFPGFVPEATLSPDIYPL 66

QY      71 SPPTRETLTKMHNPNTLYENGVCISILHPVDPOQSGELPSEKMNPTONTVRLTILSVIL 130
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       67 SPPEKRTCEMHNPNTYDGRVCLISILHAGDDPMGYESSAERNSPQSVKEKILLSWSM 126

QY      131 LNEVNTFSPANVDASVMPFKKWRDSKGKDEYAYIRKQV 169
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      127 LAEPNDSRQANVDSKMM---WRD---DREQFYKIAKQI 158

```

RESULT 20
US-09-359-967-3
; Sequence 3, Application US/09359967

1 PATENT NO. 6146624
2
3 GENERAL INFORMATION:
4
5 APPLICANT: Lal, Preeti
6
7 APPLICANT: Hillman, Jennifer L.
8
9 APPLICANT: Corley, Neil C.
10
11 TITLE OF INVENTION: HUMAN UBIQUITIN CONJUGATING ENZYMES
12
13 NUMBER OF SEQUENCES: 6
14
15 CORRESPONDENCE ADDRESS:
16
17 ADDRESSEE: Incyte Pharmaceuticals, Inc.
18
19 STREET: 3174 Porter Drive
20
21 CITY: Palo Alto
22
23 STATE: CA
24
25 COUNTRY: USA

```

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
10
```

; CLONE: 2456290
JIS-C9-359-967-3

```

Query Match      25.24;   Score 380.5;   DH 4;   Length 165;
Best local Similarity 50.34;   Prod. No. 6.90-32;
Matches 80;   Conservative 18;   Mismatches 54;   Indels 7;   Gaps 2;

QY      11  KALMLEKLSLOEPEFVEGRITLVDSDLYWHEVAITGLPNTIYEGSYFKAHLEKPPEDYPP 10
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      7  KRLMAEYKQILINPPEGIVAGPMNEEFKALINQIDCTCFQFPVPAISPPEDYPP 66
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY      71  SPPTFERLTAKMHPIINYENGWQCSILHPVDDPQSGELIPSERKNPCTNYRTILISVLS 130
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      67  SPPKMRITCEMFIPIINYDGRVCSILHAGDDPMQYESSAERPSQSVSEKLLISVWSM 126
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY      131  LNEPNTFSFANDVASVMPFKKWRDSKGDKEYAEITHKOV 169
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      127  LAEPNDSFANVDASKM---WRD-----DREDFYKIAKOL 158
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

RESULT: 21
US-08-965-689A-6
: Sequence 6, Application US/28965689A
: Patent No. 6015702
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Corlsey, Neil C.
: TITLE OF INVENTION: HUMAN BRUCHIN CONJUGATING ENZYMES
: NUMBER OF SEQUENCES: 6

```

CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 374 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,689A
FILING DATE: Hcr5with
CLASSIFICATION: 435

PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: 18-0356-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

```

Query Match 23.7%; Score 358.5; DB 3; Length 165;
Best Local Similarity 47.3%; Pred. No. 1.4e-29;
Matches 70; Conservative 26; Mismatches 49; Indels 3; Gaps 1;

```

Db 4 TAQRLKLAQLIKOSPPGIVAGPKSENNIFINCLIQGPDTPYADGVFNAKLEFPKD 63
QY 68 YPSPPTFRLTKMHPNINYEVDVCISILHPPVDQSQSLPSEKRNNTQNVRTILLSV 127
Db 64 YPLSPKLTFTPSILHNPINYEVDVCISILHSPGDDPNMYELAEERWSPVQSVKILLSV 123
QY 128 ISLLNEPNTFSPANVDASVMEKRWDSK 155
Db 124 MSHLSERPNIESGANIDACIL---WRDNK 148

RESULT 22
US-09-359-967-6
: Sequence 6, Application US/09359967
: Patent No. 6146624
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN UBQUITIN CONJUGATING ENZYMES
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/359,967
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/965,689
: FILING DATE:
: APPLICATION NUMBER: 08/933,750
: FILING DATE: September 23, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356-1 CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0535
: TELEFAX: 650-845-4166
: JTELEX:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 165 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 4257
US-09-359-967-6

Query Match 23.7%; Score 358.5; DB 4; Length 165;
Best Local Similarity 47.3%; Pred. No. 1.4e-29;
Matches 70; Conservative 26; Mismatches 49; Indels 3; Gaps 1;

QY 8 SSOKALMLEKLSLOEPVEGFRITLVDSDLYNNEVAIFGLPNTLYEGGYKKAHKKEPFD 67
Db 4 TAQRLKLAQLIKOSPPGIVAGPKSENNIFINCLIQGPDTPYADGVFNAKLEFPKD 63
QY 68 YPSPPTFRLTKMHPNINYEVDVCISILHPPVDQSQSLPSEKRNNTQNVRTILLSV 127
Db 64 YPLSPKLTFTPSILHNPINYEVDVCISILHSPGDDPNMYELAEERWSPVQSVKILLSV 123

```

```

QY 128 ISLLNEPNTFSPANVDASVMEKRWDSK 155
Db 124 MSHLSERPNIESGANIDACIL---WRDNK 148

RESULT 23
US-09-358-580-14
: Sequence 14, Application US/09358580
: Patent No. 6107545
: GENERAL INFORMATION:
: APPLICANT: Mahajan, Pramod B.
: TITLE OF INVENTION: Maize Rad6 Genes and Uses Thereof
: FILE REFERENCE: 0883
: CURRENT APPLICATION NUMBER: US/09/358,580
: CURRENT FILING DATE: 1999-07-21
: EARLIER APPLICATION NUMBER: 60/096,546
: EARLIER FILING DATE: 1998-08-14
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 14
: LENGTH: 172
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae
US-09-358-580-14

Query Match 20.5%; Score 310; DB 3; Length 172;
Best Local Similarity 30.0%; Pred. No. 1.6e-24;
Matches 66; Conservative 46; Mismatches 58; Indels 50; Gaps 4;

QY 7 TSSOKALMLEKLSLOEPVEGFRITLVDSDLYNNEVAIFGLPNTLYEGGYKKAHKKEPFD 66
Db 3 TPARKRLMRDKRKMKEDAPPCVSAPLPDNNVW WNAWILGPADIDYELQTEKLEEFDE 61
QY 67 DYPSPTFRLTKMHPNINYEVDVCISILHPPVDQSQSLPSEKRNNTQNVRTILLSV 126
Db 62 EYENKPPHVKFELSEMFHPNYANGEICDIL-----QNRWTFYDVASILTS 159
QY 127 VISLLNEPNTFSPANVDASVMEKRWDSKGGKQKRYAEIRKQVSATKALAEKQVKKVPT 186
Db 109 IOSLENDNPASNPANVEATLIF-----KHKSKYKRVKKTIVKSWEDM----- 153
QY 187 LAEYCIKTKVPSNONSDDLVDYDDSDTDDEEEED 226
Db 154 -----DDMCDQDDDDDDDDDDDDDD 372

RESULT 24
US-08-318-947A-11
: Sequence 11, Application US/08318947A
: Patent No. 5798245
: GENERAL INFORMATION:
: APPLICANT: Anderson, Paul J.
: APPLICANT: Tian, Qingsheng
: TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sughrue, Mlog, Zinn, Macpeak & Seas
: STREET: 2100 Pennsylvania Avenue, NW Suite 800
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/318,947A
: FILING DATE: 06-OCT-1994
: CLASSIFICATION: 435

```

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/133,530
 FILING DATE: 07-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Mack, Susan J.
 REGISTRATION NUMBER: 30,951
 REFERENCE/DOCKET NUMBER: A6462
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)293-7060
 TELEFAX: (202)293-2920
 TELEX: 6491103
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 172 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-318-947A-11

Query Match 20.4%; Score 308; DB 1; Length 172;
 Best Local Similarity 30.0%; Pred. No. 2.6e-24;
 Matches 66; Conservative 46; Mismatches 58; Indels 50; Gaps 4;
 QY 7 TSSOKALMLELKSIOEPVEGFRITIVDESDLYNWEVAIFGLPNTLYEGGYKAKHKEP1 66
 DB 3 TPARRLRDRKRKEDAPPGVSASPLPDNVW-KNAMIIGPADTPYEDGTFRLLLEFDE 61
 QY 67 DYPSPPTRELTMMHNPNIYENGVCISILHPVDDPQSGELPSEKRNPTQNVCTILLS 126
 DB 62 EYENKPPHVKFELSEMHPNVYANGELICDIL-----QNKWPTIYDVASILTS 108
 QY 127 VISLNEPTSPANVDASVYFRKWRDSKGKDYAEIIRKQVSAKAEKDGKGVKVTPT 186
 DB 109 IQSLFNDPASPANVEAATIF-----KHKSYKVRKVEKSWEDDM---- 153
 QY 187 LAEYCITKVPNSNDSSILLYDDIDDEDEFEEDAD 226
 DB 154 -----DDMDDDDDDDDDDDDDDD 172

RESULT 25
 US-08-795-303-11
 Sequence 11, Application US/08795303
 Patent No. 5948656
 GENERAL INFORMATION:
 APPLICANT: Anderson, Paul J.
 APPLICANT: Tian, Qingsheng
 TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
 TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sughrie, Mion, Zinn, Macpeak & Seas
 STREET: 2100 Pennsylvania Avenue, NW Suite 800
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/795,303
 FILING DATE: 04-FEB-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/318,947
 FILING DATE: 06-OCT-1994
 APPLICATION NUMBER: 08/133,530
 FILING DATE: 07-OCT-1993
 ATTORNEY/AGENT INFORMATION:

NAME: Mack, Susan J.
 REGISTRATION NUMBER: 30,951
 REFERENCE/DOCKET NUMBER: A6462
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)293-7060
 TELEFAX: (202)293-2920
 TELEX: 6491103
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 172 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-795-303-11
 Query Match 20.4%; Score 308; DB 2; Length 172;
 Best Local Similarity 30.0%; Pred. No. 2.6e-24;
 Matches 66; Conservative 46; Mismatches 58; Indels 50; Gaps 4;
 QY 7 TSSOKALMLELKSIOEPVEGFRITIVDESDLYNWEVAIFGLPNTLYEGGYKAKHKEP1 66
 DB 3 TPARRLRDRKRKEDAPPGVSASPLPDNVW-KNAMIIGPADTPYEDGTFRLLLEFDE 61
 QY 67 DYPSPPTRELTMMHNPNIYENGVCISILHPVDDPQSGELPSEKRNPTQNVCTILLS 126
 DB 62 EYENKPPHVKFELSEMHPNVYANGELICDIL-----QNKWPTIYDVASILTS 108
 QY 127 VISLNEPTSPANVDASVYFRKWRDSKGKDYAEIIRKQVSAKAEKDGKGVKVTPT 186
 DB 109 IQSLFNDPASPANVEAATIF-----KHKSYKVRKVEKSWEDDM---- 153
 QY 187 LAEYCITKVPNSNDSSILLYDDIDDEDEFEEDAD 226
 DB 154 -----DDMDDDDDDDDDDDDDDD 172

RESULT 26
 US-08-247-904B-16
 Sequence 16, Application US/08247904B
 Patent No. 5981699
 GENERAL INFORMATION:
 APPLICANT: Rolfe, Mark
 APPLICANT: Eckstein, Jens W.
 APPLICANT: Draetta, Giulio
 TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley, Hoag & Elliot
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/247,904B
 FILING DATE: 23-MAY-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MIV-029.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 832-1000
 TELEFAX: (617) 832-7000
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 152 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-247-904B-16

Query Match 19.5% Score 295; DB 2; Length 152;
Best Local Similarity 37.6% Pred. No. 4.9e-23;
Matches 65; Conservative 31; Mismatches 53; Indels 24; Gaps 4;

QY 7 TSSKALMLELSKLOEPVGHFRITLVDSN:YNHVAIFGLPNTLYHGGYFAHAKFPI 66
DB 3 TPARRRLMRDFKRLQEDPPGVGS-GAPSENNIMQWNAVIFGPGTFEGTKLVIEFSE 61
QY 67 DYPSPTPTFRELTKMMHPIYENGVDVCSISILHPPVDDPQSGELPSERWNP:QNVRIILLS 126
DB 62 EYPNKPPTVRELTKMMHPIYENVYADSGICLDIL:-----QNRKSP:YDVSSILTS 108
QY 127 VISLLNEPNTFSPANVDASVMEKRWKRSKDKKEYAEIIRKQVSATKAFAEKD 179
DB 109 IQSLDEPNPNSPANSQAALYOE:-----NKREY:-----EKRVSALVEQSWND 151

RESULT 27

US-08-767-942A-17
Sequence 17; Application US/08767942A
Patent No. 6068982

GENERAL INFORMATION:
APPLICANT: Rolfe, Mark

APPLICANT: Chiu, M. Isabel
APPLICANT: Berlin, Vivian

APPLICANT: Damagnez, Veronique
APPLICANT: Daeatta, Giulio

APPLICANT: Guillaume, Cottarel
TITLE OF INVENTION: UBIOULIN CONJUGATING ENZYMES

NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square

CITY: Boston
STATE: MA

COUNTRY: USA
ZIP: 02109-2170

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/767,942A
FILING DATE: 17-DEC-1996

ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.04

TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-767-942A-17

Query Match 19.5% Score 295; DB 3; Length 152;
Best Local Similarity 37.6% Pred. No. 4.9e-23;
Matches 65; Conservative 31; Mismatches 53; Indels 24; Gaps 4;

QY 7 TSSKALMLELSKLOEPVGHFRITLVDSN:YNHVAIFGLPNTLYHGGYFAHAKFPI 66
DB 3 TPARRRLMRDFKRLQEDPPGVGS-GAPSENNIMQWNAVIFGPGTFEGTKLVIEFSE 61

QY 67 DYPSPTPTFRELTKMMHPIYENGVDVCSISILHPPVDDPQSGELPSERWNP:QNVRIILLS 126
DB 62 EYPNKPPTVRELTKMMHPIYENVYADSGICLDIL:-----QNRKSP:YDVSSILTS 108
QY 127 VISLLNEPNTFSPANVDASVMEKRWKRSKDKKEYAEIIRKQVSATKAFAEKD 179
DB 109 IQSLDEPNPNSPANSQAALYOE:-----NKREY:-----EKRVSALVEQSWND 151

RESULT 28

US-08-318-947A-7

Sequence 7; Application US/08318947A
Patent No. 5798245

GENERAL INFORMATION:
APPLICANT: Anderson, Paul J.

APPLICANT: Tian, Qingsheng
TITLE OF INVENTION: TTA-1 BINDING PROTEINS AND ISOLATED

TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpocak & Soas

STREET: 2100 Pennsylvania Avenue, NW Suite 800
CITY: Washington

STATE: DC
COUNTRY: USA

ZIP: 20037
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,947A

FILING DATE: 06-OCT-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/133,530

FILING DATE: 07-OCT-1993
ATTORNEY/AGENT INFORMATION:

NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951

REFERENCE/DOCKET NUMBER: A6462
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060
TELEFAX: (202)293-2920

TELEX: 6491103
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-318-947A-7

Query Match 19.4% Score 293; DB 1; Length 152;

Best Local Similarity 37.6% Pred. No. 8e-23;
Matches 65; Conservative 31; Mismatches 53; Indels 24; Gaps 4;

QY 7 TSSKALMLELSKLOEPVGHFRITLVDSN:YNHVAIFGLPNTLYHGGYFAHAKFPI 66
DB 3 TPARRRLMRDFKRLQEDPPGVGS-GAPSENNIMQWNAVIFGPGTFEGTKLVIEFSE 61
QY 67 DYPSPTPTFRELTKMMHPIYENGVDVCSISILHPPVDDPQSGELPSERWNP:QNVRIILLS 126
DB 62 EYPNKPPTVRELTKMMHPIYENVYADSGICLDIL:-----QNRKSP:YDVSSILTS 108

QY 127 VISLLNEPNTFSPANVDASVMEKRWKRSKDKKEYAEIIRKQVSATKAFAEKD 179
DB 109 IQSLDEPNPNSPANSQAALYOE:-----NKREY:-----EKRVSALVEQSWND 151

RESULT 29

MEDIUM TYPE:

```

? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/795,303
? FILING DATE: 04-FEB-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/318,947
? FILING DATE: 06-OCT-1994
? APPLICATION NUMBER: 08/133,530
? FILING DATE: 07-OCT-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Mack, Susan J.
? REGISTRATION NUMBER: 30,952
? REFERENCE/DOCKET NUMBER: A6462
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202)293-7060
? TELEFAX: (202)293-2920
? TELEX: 6491103
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 152 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-795-301-6

Query Match 19.3%; Score 291; DR 2; Length 152;
Best Local Similarity 37.6%; Pred. No. 1.3e-22;
Matches 65; Conservative 30; Mismatches 54; Indels 24; Gaps 4;

QY 7 ISSQALMLKLSGEEPEVEGRITLVDESDLYNNWEVAIFGLPNTLYEGYFKAHKFP 66
DB 3 IPARRRLMRDEKRIQEDPPAGVS-GAPENNIMVNNAVIEGCTPFEGCTFKLIEFSE 61

QY 67 DYPSPTPTFRELKMWHPNIYENGVCISILHPVDDPQSGELPSERNWPTONVATILS 126
DB 62 EYPNKPPTVRELKMFHPNYADGSGICIDIL-----QNRWSTYDVSSILTS 109

QY 127 VISLNEPNTSPANVDASVMPFKWDSKGGKIKYAFITRKQVSATKAAEKD 179
DB 109 IQSLIDENPNPANSQAQIYOE-----NKREY-----EKRVSAIVGQSWND 151

RESULT 32
US-08-318-947A-8
? Sequence 8, Application US/08318947A
? Patent No. 5798245
? GENERAL INFORMATION:
? APPLICANT: Anderson, Paul J.
? APPLICANT: Tian, Qingsheng
? TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
? TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
? NUMBER OF SEQUENCES: 21
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Sughrue, Miron, Zinn, Macpeak & Seas
? STREET: 2100 Pennsylvania Avenue, NW Suite 800
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/318,947A
? FILING DATE: 06-OCT-1994
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:

```

```

? APPLICATION NUMBER: 08/133,530
? FILING DATE: 07-OCT-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Mack, Susan J.
? REGISTRATION NUMBER: 30,952
? REFERENCE/DOCKET NUMBER: A6462
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202)293-7060
? TELEFAX: (202)293-2920
? TELEX: 6491103
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 152 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-318-947A-8

Query Match 18.9%; Score 286; DR 1; Length 152;
Best Local Similarity 36.4%; Pred. No. 4.3e-22;
Matches 63; Conservative 33; Mismatches 53; Indels 24; Gaps 4;

QY 7 ISSQALMLKLSGEEPEVEGRITLVDESDLYNNWEVAIFGLPNTLYEGYFKAHKFP 66
DB 3 IPARRRLMRDEKRIQEDPPAGVS-GAPENNIMVNNAVIEGCTPFEGCTFKLIEFSE 61

QY 67 DYPSPTPTFRELKMWHPNIYENGVCISILHPVDDPQSGELPSERNWPTONVATILS 126
DB 62 EYPNKPPTVRELKMFHPNYADGSGICIDIL-----QNRWSTYDVSSILTS 109

QY 127 VISLNEPNTSPANVDASVMPFKWDSKGGKIKYAFITRKQVSATKAAEKD 179
DB 109 IQSLIDENPNPANSQAQIYOE-----NKREY-----EKRVSAIVGQSWND 151

RESULT 33
US-08-795-303-8
? Sequence 8, Application US/08795303
? Patent No. 5948656
? GENERAL INFORMATION:
? APPLICANT: Anderson, Paul J.
? APPLICANT: Tian, Qingsheng
? TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
? TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
? NUMBER OF SEQUENCES: 21
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Sughrue, Miron, Zinn, Macpeak & Seas
? STREET: 2100 Pennsylvania Avenue, NW Suite 800
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/795,303
? FILING DATE: 04-FEB-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/318,947
? FILING DATE: 06-OCT-1994
? APPLICATION NUMBER: 08/133,530
? FILING DATE: 07-OCT-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Mack, Susan J.
? REGISTRATION NUMBER: 30,952
? REFERENCE/DOCKET NUMBER: A6462
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202)293-7060

```

```

; TELEFAX: (202)293-2920
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-795-303-8

Query Match      18.9%; Score 286; DB 2; Length 152;
Best Local Similarity 36.4%; Pred. No. 4,3e-22;
Matches 63; Conservative 33; Mismatches 53; Indels 24; Gaps 4;

QY 7 TSSQKALMELKSLQEEPRVEGFRITLVDSGLYNNMFVAIFGLPNTLYENGYKAIHKFP 66
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3 TTARRRLMRDFKRMQDPAGVSASPSVSNVNL-KNAVILGPADTPFEFGTKLVLSFCE 61
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 67 DYPYSPPTFRFLTKMHPNIYENGDCVCSILHPVDDPQSGELFSRWNPNTONRTILLS 126
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 62 EYENKPTVRSKMPHPNYADSCICIDL-----QNRWSPYDVSVILTS 108
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 127 VISLNEPNTSPANDASVYFKWRDSKGDKYAEIIRKQV 179
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 109 IQSLLEPNPNSPANQAOLYOE-----NKREY---EKRVSAIVEQSWRD 151
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 34
US-08-318-947A-10
; Sequence 10. Application US/08318947A
; Patent No. 5798245
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; APPLICANT: Tian, Qingsheng
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,947A
; FILING DATE: 06-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/133,530
; FILING DATE: 07-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,951
; REFERENCE/DOCKET NUMBER: A6462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-2920
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-318-947A-10

```

```

Query Match      18.8%; Score 284; DB 1; Length 151;
Best Local Similarity 35.8%; Pred. No. 6,8e-22;
Matches 58; Conservative 32; Mismatches 53; Indels 20; Gaps 3;

QY 7 TSSQKALMELKSLQEEPRVEGFRITLVDSGLYNNMFVAIFGLPNTLYENGYKAIHKFP 66
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3 TTARRRLMRDFKRMQDPAGVSASPSVSNVNL-KNAVILGPADTPFEFGTKLVLSFCE 61
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 67 DYPYSPPTFRFLTKMHPNIYENGDCVCSILHPVDDPQSGELFSRWNPNTONRTILLS 126
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 62 QYENKPTLVKFS-MFHPNYANGELCDIL-----QNRWSPYVAALTS 108
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 127 VISLNEPNTSPANDASVYFKWRDSKGDKYAEIIRKQV 169
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 109 IQSLLEPNPNSPANQAALHRE-----NKKEYVRKVRKTV 145
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 35
US-08-795-303-10
; Sequence 10. Application US/08795303
; Patent No. 5948656
; GENERAL INFORMATION:
; APPLICANT: Anderson, Paul J.
; APPLICANT: Tian, Qingsheng
; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
; TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, NW Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,303
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/318,947
; FILING DATE: 06-OCT-1994
; APPLICATION NUMBER: 08/233,530
; FILING DATE: 07-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mack, Susan J.
; REGISTRATION NUMBER: 30,951
; REFERENCE/DOCKET NUMBER: A6462
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-2920
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-795-303-10

Query Match      18.8%; Score 284; DB 2; Length 151;
Best Local Similarity 35.8%; Pred. No. 6,8e-22;
Matches 58; Conservative 32; Mismatches 53; Indels 20; Gaps 3;

QY 7 TSSQKALMELKSLQEEPRVEGFRITLVDSGLYNNMFVAIFGLPNTLYENGYKAIHKFP 66
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 3 TTARRRLMRDFKRMQDPAGVSASPSVSNVNL-KNAVILGPADTPFEFGTKLVLSFCE 61
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```


Sequence 9, Application US/08318947A
Patent No. 5798245

GENERAL INFORMATION:

APPLICANT: Anderson, Paul J.
APPLICANT: Tian, Qingsheng
TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, NW Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,947A
FILING DATE: 06-OCT-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/133,530
FILING DATE: 07-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951
REFERENCE/DOCKET NUMBER: A6462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-2920
TELEX: 6491103

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-318-947A-9

Query Match 17.9%; Score 271; DB 1; Length 151;

Best Local Similarity 35.6%; Pred. No. 1.5e-20;

Matches 57; Conservative 34; Mismatches 49; Indels 20; Gaps 3;

QY 7 TSSQKALMELKSLQEPVEGFRITLVDSLYNVEVAIFGLPNTIYEGGYFAHIFPI 66

Db 3 TPARRRLMRDFKRLQEDPPTGVSGAPTD-NNIMTNNAVIFGPHDTPFDGTFKLTIEFE 61

QY 67 DYFSPPTFERFLTKMHPNIYENGDVGISILHPPVDDPQSGELPSENNPTQNVRTILLS 126

Db 62 EYPNKPPTVFEVSKVHPNVYADGGTICLDIL-----QNRWSPRYDVSAILTS 108

QY 127 VISLLNEPNTFSPANVDASVMFKWRDSKKGKEYAEIIR 166

Db 109 IQSLSDPNPNSPANSTAAQLYKENR-----REYKRVK 142

RESULT 40

US-08-795-303-9

Sequence 9, Application US/08795303

Patent No. 5948656

GENERAL INFORMATION:

APPLICANT: Anderson, Paul J.
APPLICANT: Tian, Qingsheng
TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
TITLE OF INVENTION: COMPLEMENTARY DNA ENCODING THE SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas

STREET: 2100 Pennsylvania Avenue, NW Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,303
FILING DATE: 04-FEB-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/318,947
FILING DATE: 06-OCT-1994
APPLICATION NUMBER: 08/133,530
FILING DATE: 07-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Mack, Susan J.
REGISTRATION NUMBER: 30,951
REFERENCE/DOCKET NUMBER: A6462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-2920
TELEX: 6491103

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-795-303-9

Query Match 17.9%; Score 271; DB 2; Length 151;

Best Local Similarity 35.6%; Pred. No. 1.5e-20;

Matches 57; Conservative 34; Mismatches 49; Indels 20; Gaps 3;

QY 7 TSSQKALMELKSLQEPVEGFRITLVDSLYNVEVAIFGLPNTIYEGGYFAHIFPI 66

Db 3 TPARRRLMRDFKRLQEDPPTGVSGAPTD-NNIMTNNAVIFGPHDTPFDGTFKLTIEFE 61

QY 67 DYFSPPTFERFLTKMHPNIYENGDVGISILHPPVDDPQSGELPSENNPTQNVRTILLS 126

Db 62 EYPNKPPTVFEVSKVHPNVYADGGTICLDIL-----QNRWSPRYDVSAILTS 108

QY 127 VISLLNEPNTFSPANVDASVMFKWRDSKKGKEYAEIIR 166

Db 109 IQSLSDPNPNSPANSTAAQLYKENR-----REYKRVK 142

Search completed: April 10, 2003, 10:46:58

Job time : 17.2612 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 10, 2003, 10:29:19 ; Search time 44.7882 Seconds
(without alignments)
1297.334 Million cell updates/sec

Title: US-09-930-026-2

Perfect score: 1511

Sequence: 1 MAQQQWTSKALMIFIKSL.....AMAPQOKPIHSGWGNTHSSC 282

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 135 summaries

Database :

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organella.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vvertebrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriap.*
- 17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1235	81.7	238	11 Q8VDE5	Q8VDE5 mus musculus
2	1232	81.5	238	4 Q9N64	Q9N64 homo sapien
3	870	57.6	341	5 Q9VUR4	Q9VUR4 drosophila
4	757.5	50.1	329	5 Q9VXN7	Q9VXN7 caenorhabdi
5	643	42.6	270	5 Q96Q05	Q96Q05 drosophila
6	516	34.1	235	5 Q76542	Q76542 dictyosteli
7	454.5	30.1	168	5 Q8VY63	Q8VY63 drosophila
8	452	29.9	169	10 Q943L1	Q943L1 oryza sativ
9	450	29.8	166	10 Q42540	Q42540 arabidopsis
10	450	29.8	187	10 Q8RXQ3	Q8RXQ3 arabidopsis
11	444	29.4	169	10 Q22656	Q22656 zea mays (m
12	444	29.4	169	10 Q24560	Q24560 zea mays (m
13	439	29.1	166	10 Q42541	Q42541 arabidopsis
14	434.5	28.8	168	5 Q9VFE8	Q9VFE8 drosophila
15	417	27.6	167	3 Q9YB18	Q9YB18 schizosacch
16	398	26.3	166	3 Q9C2A5	Q9C2A5 neurospora

17	389	25.7	167	5 Q9VXER	Q9VXER drosophila
18	385.5	25.5	172	5 Q8SS54	Q8SS54 ercephalito
19	382	25.3	200	5 Q9N9Z5	Q9N9Z5 drosophila
20	381	25.2	170	5 Q9U102	Q9U102 caenorhabdi
21	373.5	24.7	257	3 Q8X0N3	Q8X0N3 neurospora
22	359.5	23.8	148	11 Q9QX59	Q9QX59 mus musculu
23	317.5	21.0	136	11 Q9CZJ4	Q9CZJ4 mus musculu
24	304.5	20.2	216	5 Q8SSK3	Q8SSK3 encephalito
25	304	20.1	162	5 Q8SR17	Q8SR17 encephalito
26	294	19.5	152	11 Q9DDJ6	Q9DDJ6 mus musculu
27	293	19.4	152	4 Q96FX4	Q96FX4 homo sapien
28	293	19.4	152	11 Q9Z255	Q9Z255 mus musculu
29	293	19.4	152	13 Q9W6F3	Q9W6F3 galles gall
30	286	18.9	151	3 Q96UP5	Q96UP5 emeticolia
31	283	18.7	151	3 P78717	P78717 acetria aae
32	280	18.5	152	10 Q9AVP0	Q9AVP0 nicotiana t
33	277	18.3	152	10 Q9MVR9	Q9MVR9 avicennia m
34	276	18.3	152	10 Q9AVN9	Q9AVN9 nicotiana t
35	274	18.1	124	6 Q95L52	Q95L52 bos taurus
36	270	17.9	148	5 Q9GNC1	Q9GNC1 leishmania
37	269	17.8	124	6 Q9N1X5	Q9N1X5 bos taurus
38	269	17.8	126	6 Q9N1X6	Q9N1X6 epus cabal
39	269	17.8	146	20 Q94AD2	Q94AD2 arabidopsis
40	269	17.8	152	10 Q8SBC1	Q8SBC1 oryza sativ
41	267.5	17.7	186	5 Q9U5Q3	Q9U5Q3 entodinium
42	266	17.6	126	6 Q9KX7	Q9KX7 sus scrofa
43	262	17.3	97	10 Q8VWC2	Q8VWC2 narcissus p
44	261	17.3	155	3 Q9UTN8	Q9UTN8 schizosacch
45	259	17.1	148	10 Q48555	Q48555 zea mays (m
46	255.5	16.9	148	10 Q9SVF9	Q9SVF9 mesembryant
47	254	16.8	204	5 Q8SSX8	Q8SSX8 encephalito
48	252.5	16.7	179	11 Q9N1C1	Q9N1C1 mus musculu
49	252	16.7	148	10 Q42897	Q42897 lycopersico
50	252	16.7	178	10 Q9FFH9	Q9FFH9 arabidopsis
51	252	16.7	374	10 Q94DA8	Q94DA8 oryza sativ
52	251	16.6	148	10 Q94F47	Q94F47 arabidopsis
53	250	16.5	148	10 Q43821	Q43821 pisum sativ
54	248	16.4	146	10 Q9SH72	Q9SH72 arabidopsis
55	248	16.4	147	10 Q8XAO	Q8XAO oryza sativ
56	247	16.3	147	10 Q8S920	Q8S920 oryza sativ
57	246.5	16.3	140	4 Q9BQP1	Q9BQP1 homo sapien
58	245	16.2	158	5 Q8SR36	Q8SR36 encephalito
59	243	16.1	148	10 Q8S919	Q8S919 oryza sativ
60	243	16.1	148	10 Q8RXQ0	Q8RXQ0 arabidopsis
61	242	16.0	140	10 Q8VWV2	Q8VWV2 pinus pinas
62	240	15.9	160	5 Q77397	Q77397 plasmodium
63	239	15.8	178	5 Q9VTF6	Q9VTF6 drosophila
64	237.5	15.7	172	13 Q9PVW3	Q9PVW3 carassius a
65	237	15.7	148	10 Q9SFE4	Q9SFE4 arabidopsis
66	236.5	15.7	158	13 Q9W6H5	Q9W6H5 brachydanio
67	236	15.6	148	10 Q9ZT28	Q9ZT28 pinus resin
68	236	15.6	149	10 Q9C9Y7	Q9C9Y7 arabidopsis
69	235.5	15.6	157	13 Q9DDJ0	Q9DDJ0 brachydanio
70	235	15.6	184	4 Q9BQZ5	Q9BQZ5 homo sapien
71	234.5	15.5	354	5 Q9J7T6	Q9J7T6 drosophila
72	233.5	15.5	147	11 Q9D7F5	Q9D7F5 mus musculu
73	233.5	15.5	153	10 Q9XHP3	Q9XHP3 catharanthu
74	233	15.4	147	4 Q9Y2X8	Q9Y2X8 n. n. n. n. n.
75	233	15.4	154	10 Q9C8X7	Q9C8X7 arabidopsis
76	232.5	15.4	201	5 Q9NKC1	Q9NKC1 drosophila
77	231.5	15.3	197	4 Q9NP08	Q9NP08 homo sapien
78	231	15.3	147	3 Q9UV92	Q9UV92 magrapothe
79	231	15.3	147	10 Q987H6	Q987H6 quillardia
80	230	15.2	148	10 Q9FKT3	Q9FKT3 arabidopsis
81	228	15.1	76	4 Q9BUR6	Q9BUR6 homo sapien
82	228	15.1	147	3 Q74196	Q74196 collietritic
83	227.5	15.1	166	5 Q95017	Q95017 caenorhabdi
84	227.5	15.1	191	10 Q9L1Y1	Q9L1Y1 oryza sativ
85	226.5	15.0	152	11 Q9DAJ6	Q9DAJ6 mus musculu
86	225.5	14.9	148	10 Q42973	Q42973 oryza sativ
87	225.5	14.9	204	11 Q9CQ37	Q9CQ37 mus musculu
88	225	14.9	153	5 Q21633	Q21633 caenorhabdi
89	224.5	14.9	152	11 Q9EQX9	Q9EQX9 rat.t.s norv

09/ 920 026

90 224.5 14.9 159 5 062622 C62622 drosophila
 91 223.5 14.8 195 10 09C604 Q9C604 arabidopsis
 92 223 14.8 133 11 063546 Q63546 ratcus norv
 93 222.5 14.7 153 10 09F248 Q9F248 arabidopsis
 94 222 14.7 183 10 09LRS2 Q9LRS2 arabidopsis
 95 221 14.6 151 5 09VTS5 Q9VTS5 drosophila
 96 220 14.6 147 11 09DJS1 Q9DJS1 mus musculu
 97 219.5 14.5 153 10 08W011 Q8W011 oryza sativ
 98 218 14.4 234 5 09VY83 Q9VY83 drosophila
 99 217.5 14.4 175 5 08SOR1 Q8SOR1 encephalito
 100 216.5 14.3 118 11 09D6V0 Q9D6V0 mus musculu
 101 216.5 14.3 153 10 094A97 Q94A97 arabidopsis
 102 215.5 14.3 148 3 09J685 Q9J685 schizosacch
 103 213.5 14.1 118 4 096RP6 Q96RP6 homo sapien
 104 213.5 14.1 154 5 096840 Q96840 drosophila
 105 212.5 14.1 185 5 095XX0 Q95XX0 caenorhabd
 106 211 14.0 118 11 09D808 Q9D808 mus musculu
 107 211 14.0 251 10 09FF66 Q9FF66 arabidopsis
 108 210.5 13.9 164 5 09VQ00 Q9VQ00 drosophila
 109 210.5 13.9 181 10 09LJ25 Q9LJ25 arabidopsis
 110 210.5 13.9 190 5 09VX25 Q9VX25 drosophila
 111 209 13.8 163 10 09ZVA6 Q9ZVA6 arabidopsis
 112 207.5 13.7 195 10 08S1Y5 Q8S1Y5 oryza sativ
 113 206 13.6 148 5 09NEB4 Q9NEB4 leishmania
 114 205.5 13.6 159 5 09NGR4 Q9NGR4 dictyosteli
 115 205 13.6 151 4 09HAG6 Q9HAG6 homo sapien
 116 204.5 13.5 491 5 09V165 Q9V165 drosophila
 117 204 13.5 147 4 09NTT1 Q9NTT1 homo sapien
 118 204 13.5 151 4 096R02 Q96R02 homo sapien
 119 204 13.5 151 11 08VDM4 Q8VDM4 mus musculu
 120 202.5 13.4 147 8 098S78 Q98S78 quillardia
 121 202 13.4 217 3 074810 Q74810 schizosacch
 122 202 13.4 409 10 09LV54 Q9LV54 arabidopsis
 123 202 13.4 409 10 094186 Q94186 arabidopsis
 124 201.5 13.3 160 10 08W2X7 Q8W2X7 oryza sativ
 125 201.5 13.3 180 5 09VZ73 Q9VZ73 drosophila
 126 201 13.3 98 13 098T47 Q98T47 platichtyus
 127 201 13.3 199 5 08T026 Q8T026 bombyx mori
 128 200.5 13.3 223 11 0921J4 Q921J4 mus musculu
 129 200 13.2 194 10 043780 Q43780 lycopersico
 130 199.5 13.2 148 5 094490 Q94490 dictyosteli
 131 199.5 13.2 425 5 095039 Q95039 paramecium
 132 198.5 13.1 80 4 09WXB3 Q9WXB3 homo sapien
 133 198.5 13.1 81 11 08VH44 Q8VH44 mus musculu
 134 198.5 13.1 201 4 0961R5 Q961R5 homo sapien
 135 198.5 13.1 201 11 091W82 Q91W82 mus musculu

ALIGNMENTS

RESULT 1
 Q8VDE5 PRELIMINARY: PRT: 238 AA.
 AC Q8VDE5
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Ubiquitin conjugating enzyme 3b.
 GN 1200003MILRIK OR UBC3B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Semplici F., Oliviero S.,
 RT "Cloning a new Ubc3 like gene (Ubc3b)."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ240086; CACh0335.1;
 DR MGI: 1914865; 1200003MILRIK.
 DR InterPro: IPR000608; UBU_conjugat.
 DR Pfam: PF00179; UQ_conj_1.

DR ProDom: P0000461; UBU_conj_dat: 1.
 DR SMART: SM00212; UBCc; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; UNKNOWN_1.
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2;
 SQ SEQUENCE 238 AA: 27111 MW: 890A93776C36566 CRC64;
 Query Match 81.7%; Score 1235; D5.1; Length 238;
 Best Local Similarity 99.1%; Pred. No. 3e-103;
 Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAQQMTSSOKALMELEK:KSQEPVFGFRITLVDSGLYNKFWVAFLPMLYFGGYFKA 60
 Db 1 MAQQMTSSOKALMELEK:KSQEPVFGFRITLVDSGLYNKFWVAFLPMLYFGGYFKA 60
 QY 61 HKFPIDYSPPIFREL:KMHNPNTYENGWCISILHPVDDPQSGILPSEKWNPTNV 120
 Db 61 HKFPIDYSPPIFREL:KMHNPNTYENGWCISILHPVDDPQSGILPSEKWNPTNV 120
 QY 121 RTLLSVLSLNEPTSPANVCASVWFKWDKSGKIKYAEIIRKQVSNATKAEAKG 180
 Db 121 RTLLSVLSLNEPTSPANVCASVWFKWDKSGKIKYAEIIRKQVSNATKAEAKG 180
 QY 181 VKVPTLAECYCIKTKVPSNDSSDLIYDLYGDDIDRDEREDADYDNDSSG 234
 Db 181 VKVPTLAECYCIKTKVPSNDSSDLIYDLYGDDIDRDEREDADYDNDSSG 234
 RESULT 2
 Q9NX64 PRELIMINARY: PRT: 216 AA.
 AC Q9NX64
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 26, Last annotation update)
 DE CDNA FLJ20419 f1s, clone KA702435.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Katanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isoqali T., Sugano S.,
 RT "NEO human cDNA sequencing project."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AK000426; BAA9156.1;
 DR HSP: Q02159; 2BCZ.
 DR InterPro: IPR000608; UBU_conjugat.
 DR Pfam: PF00179; UQ_conj_1.
 DR ProDom: P0000461; UBU_conjugat: 1.
 DR SMART: SM00212; UBCc; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1;
 DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2;
 KW Ligase; Ubiquitin conjugation.
 SQ SEQUENCE 238 AA: 27112 MW: F071C7A231B4E451 CRC64;

Query Match 81.5%; Score 1242; DB 4; Length 238;
 Best Local Similarity 98.7%; Pred. No. 3e-103;
 Matches 231; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAQQMTSSOKALMELEK:QEPVFGFRITLVDSGLYNKFWVAFLPMLYFGGYFKA 60
 Db 1 MAQQMTSSOKALMELEK:QEPVFGFRITLVDSGLYNKFWVAFLPMLYFGGYFKA 60

```

QY 61 HIKPDPYSPPTFRFLTKMHPNIYENGVCISILHPPVDPQSGELPSERNPTQNV 120
DB 61 HIKPDPYSPPTFRFLTKMHPNIYENGVCISILHPPVDPQSGELPSERNPTQNV 120
QY 121 KTLSSVLSLNPNTFSFANYDASVMFKWRDCKGKDEYAFIIRKQVSA:KAPAEKDG 180
DB 121 KTLSSVLSLNPNTFSFANYDASVMFKWRDCKGKDEYAFIIRKQVSA:KAPAEKDG 180
QY 181 VKVPTTLIARYCIKTKVPSNDSSLLYDDLDIDDEFEEDADCYDDDDSG 234
DB 181 VKVPTTLIARYCIKTKVPSNDSSLLYDDLDIDDEFEEDADCYDDDDSG 234
RESULT 3
QYVUR4 PRELIMINARY; PRI: 341 AA.
AC Q9VUR4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE CG7656 protein.
GN CG7656.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Schorer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.P., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Bernan R.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeqwan C.,
RA Jaisi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang P., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang P., Lai Z.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith I.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venier E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT *The genome sequence of Drosophila melanogaster.*;
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.

```

```

CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR PMBL: AE003530; AAF49611.1; -.
DR HSSP: Q02159; 20C2.
DR FlyBase: FBgn0036516; UQ-conjucat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: P500046; UQ-conjugat; 1.
DR SMART: SM00212; UBCc; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 341 AA; 37655 MW; 8F1EC9D39E62F429 CRC64;
Query Match 57.9%; Score 879; DH 5; length 341;
Best Local Similarity 58.7%; Pred. No. 2,4e-76;
Matches 169; Conservative 42; Mismatches 53; Indels 24; Gaps 7;
QY 7 TSSAKALMLELKSLOEPEVGFRLTVLDSHLYNFWVAIPLPNTLYFGGYFKAHKEPH 66
DB 61 SSAVALAHYKSGEEPVEGRVKLINDFMFWALEFPPTILYQGYFKAHKEPH 120
QY 67 DYPYSPPTFRFLTKMHPNIYENGVCISILHPPVDPQSGELPSERNPTQNVKTLSS 126
DB 121 DYPYSPPTFRFLTKMHPNIYENGVCISILHPPVDPQSGELPSERNPTQNVKTLSS 180
QY 127 VISLNPNTFSFANYDASVMFKWRDCKGKDEYAFIIRKQVSA:KAPAEKDGKVFVT 186
DB 181 VISLNPNTFSFANYDASVMFKWRDCKGKDEYAFIIRKQVSA:KAPAEKDGKVFVT 240
QY 187 LAEYC:K-TKVPNSNDSSLLYDDLDIDDEFEEDADCYDDDDSG 234
DB 241 LKDYCLKPKTKPTESIGIANFYDDDFDLEEDIPSDGDFDEDDDDGEEH--KDELSA 238
QY 235 ----MRSRDVLIQCPATIPSAKGRGASGLAMAPCKTIHSGWNT 278
DB 299 TAPISKNGSGSKCKNGLVRAAAGAD--IAESADG-----SKGCEI 343
RESULT 4
QYVUR4 PRELIMINARY; PRI: 329 AA.
AC Q9VUR4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE Hypothetical 37.4 kDa protein.
GN Y71G12B.15
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT *Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.*;
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Du H., Lamar B., Wohldmann P.;
RT *The sequence of C. elegans cosmid Y71G12B.*;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT *Direct Submission.*;

```



```

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN CONJUGATING ENZYME FAMILY.
DR EMBL; AC025726; AAK73914.2; -.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UBQ_con; 1.
DR ProDom: P0000461; UBQ_conjugat; 1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; UNKNOWN.1.
DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW Hypothetical protein; Ligase; Ubiquitin conjugation.
SQ SEQUENCE 329 AA; 37378 MW; 553C94F673B2F109 CRC64;

Query Match 50.1%; Score 757.5; DB 5; Length 329;
Best Local Similarity 60.1%; Pred. No. 3.3e-60;
Matches 143; Conservative 41; Mismatches 45; Indels 9; Gaps 5;

QY 1 MAQDQWISS--OKALMLELKSLOEERVEGFRITIVDESILYNNVEVAIG;PNTLYEGGYF 58
Db 87 MDSKASTSGALRALTWELKALQSQVEGTID-VNEDNLFVMTVGIYGPPTLYGGYF 145
QY 59 KAHIKFPIDYSPPTFRFTJTKMHPNIYENGDCISTLHPVPDDPOSGELPSEKKNPTQ 118
Db 146 KASTRPSNYPSPSPSKFTTKVHPNHYVNGDLCSILHSPIDDDPOSGELPSEKKNPTQ 205
QY 119 NVRTILLSVISLNEPNTFSPANVDASVMFRKWRDSKGKQKAEYAEIRKQVSATKAERK 178
Db 206 SVRTILLSVISLNEPNTFSPANVDASVMFRKWRDS--QDEYAKIVTKQVSKKVAQK 263
QY 179 DGVKVPYTTLAECYTKTKVPSON--SDLLYDDIDYDDIDDEDEHFDAC---YUDD 232
Db 264 DGIQVPETIEECYKWAPOQDDVDIDYNDDEFCQYDDDEDEDEDECGSDYND 321

RESULT 5
Q96005 PRELIMINARY; PRT; 270 AA.
AC Q96005;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE LD40324p.
GN CG7856.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA Champé M., Chavez C., Dorsett V., Farfan B., Frise E., George K.,
RA Gonzalez M., Guarin H., Li P., Iiao G., Miranda A., Mungall C.J.,
RA Nuno J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN CONJUGATING ENZYME FAMILY.
DR EMBL; AY051915; AAK93339.1; -.
DR FlyBase; FBgn0036516; CG7656.
DR InterPro: IPR000608; UBQ_conjugat.

```

```

DR Pfam: PF00179; UBQ_con; 1.
DR ProDom: P0000461; UBQ_conjugat; 1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; UNKNOWN.1.
DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 270 AA; 29223 MW; 2536F7EAC4016CB CRC14;

Query Match 42.6%; Score 643; DB 5; Length 270;
Best Local Similarity 56.2%; Pred. No. 5.4e-50;
Matches 131; Conservative 30; Mismatches 46; Indels 26; Gaps 8;

QY 62 IKFFIDYSPPTFRFTJTKMHPNIYENGDCISTLHPVPDDPOSGELPSEKKNPTQNV 122
Db 47 MKFPHDYSPSPISRLTKVHPNHYVNGDLCSILHSPIDDDPOSGELPSEKKNPTQNV 106
QY 122 TILLSVISLNEPNTFSPANVDASVMFRKWRDSKGKQKAEYAEIRKQVSATKAERKQV 181
Db 107 TILLSVISLNEPNTFSPANVDASVMFRKWRDS--QDNEYPNIRKQALANAFAPKEPT 164
QY 182 KVPYTTLAECYTKTKVPSONSSD--LYSDLY-----DQIDSGEUEEEDADQYD 229
Db 165 VPMILJEDYCIAPFKPTTESGLDANFYDDPLETDDGLPSDDCFDEHDDSDSD-- 222
QY 230 DDDSG---MRSRDVLLQCPCTALPSGAKGASGDLAMAPQVGRFHSWGNT 278
Db 223 DEDSATAPISKNNKSSCKNGLVREANAGAD-DATFSDLD-----SGKSET 265

RESULT 6
076542 PRELIMINARY; PRT; 235 AA.
AC 076542;
DT 01-NOV-1998 (TrEMBLrel. 68, Created)
DT 01-NOV-1998 (TrEMBLrel. 68, last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE Ubiquitin-conjugating enzyme protein UbcC.
GN UBC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA Tranfar N., Loomis W.F.
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN CONJUGATING ENZYME FAMILY.
DR EMBL; AF076597; AAC27763.1; -.
DR HSSP; Q02559; 2UCZ.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UBQ_con; 1.
DR ProDom: P0000461; UBQ_conjugat; 1.
DR SMART; SM00212; UBC.1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 235 AA; 27223 MW; F4914E09F759700 CRC64;

Query Match 34.1%; Score 524; DB 5; Length 235;
Best Local Similarity 44.3%; Pred. No. 1.3e-38;
Matches 112; Conservative 32; Mismatches 67; Indels 42; Gaps 6;

QY 8 SSKAKMLELKSLOEERVEGFRITIVDESILYNNVEVAIG;PNTLYEGGYKHAHKPTD 67
Db 2 SAKLIDNQPKKISSEPIEGVAFELVDE-NLEFWRAYLGGPPTQVGGIFQ;GNKPFND 60
QY 68 YVSPPTFRFTJTKMHPNIYENGDCISTLHPVPDDPOSGELPSEKKNPTQNVRTIC 127

```

Db 61 YPMSPVLLVQSEWHPNRYVYKGVKVCISILHPPGDETSGLPEERWLPVTIVQTIISV 120
 QY 128 ISLSENPNTFSANVDASVMEKRWDSKGDKEYAEIIRKQVSAIKAKAPKDGKVKVPTTL 187
 DB 121 ISLSAINTSSPANDVASVEKNNRDN-----YKRIREIV-----QKANKLVPSHI 167
 QY 188 AXCICKTKVPSNDSS-----DLVYDLYD-DDIDDEDEEDAD 226
 DB 168 -----KIHPTDPIVERAKVEKILINKPMQFYDDYDDNDYDYDYRENDDEED 220
 QY 227 CYDNDGSGHRSR 239
 DB 221 YNDDDEGIDEED 233

RESULT 7
 Q8SYG3 PRELIMINARY: PRI: 168 AA.

AC Q8SYG3;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE REG3412p
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Paradas V., Park S.,
 RA Patel S., Phanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celisner S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DDJ databases.
 DR EMBL: AY071574; AAL49196.1;
 SQ SEQUENCE 168 AA; 19366 MW; 8F3A55FE058A4BC8 CRC64;

Query Match 30.18; Score 454.5; DB 5; Length 168;
 Best Local Similarity 50.38; Pred. No. 3e-43;
 Matches 86; Conservative 28; Mismatches 48; Indels 11; Gaps 3;

QY 6 MTSOKALML--ELKSIQEPVFCFRITLVDSILYNNHVAIFGLPNLYEGSYFKAHIK 63
 DB 1 MSLQSLKKQLKALNKNVFGSAGLIDEND:FRMEVLIGPPDTIYEGSFFKAHLY 60
 QY 64 FPIDYPSPTFRFLTKMHIPNIYENGVCISILHPPVDDFGSGLFSPERNPTQNVRTI 123
 DB 61 FPKEYPLEPRMKFVTEIWHIPNLKNGVCISILHPPGDDKNGYFKASERWLPVHIVE: 120
 QY 124 LLSVISLNEPNTFSANVDASVMEKRWDSKGDKEYAEIIRKQVSAIKAEAK 178
 DB 121 LLSVISLADPNDESANVDAA---KEWRSE-----VTEKRVKVARCKRSGEE 166

RESULT 8
 Q943LI PRELIMINARY: PRI: 169 AA.

AC Q943LI;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Putative ubiquitin carrier protein UBC7.
 GN P003131.24.
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPUNBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RI "Oryza sativa nipunbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0031b11.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC 1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC 1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC 1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC 1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC 1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AP032331; BAB67890.1;
 DR InterPro: IPR006068; UQC_conjugat.
 DR Pfam: PF00179; UQC_con: 1.
 DR ProDom: PDC00461; UQC_conjugat: 1.
 DR PROSITE: PS00181; UBIQUITIN_CONJUGAT_1; UNKNOWN_1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ligase; Ubiquitin conjugation.
 SQ SEQUENCE 169 AA; 19013 MW; 9F9F310DECF8469F CRC64;

Query Match 29.98; Score 452; DB 5; Length 163;
 Best Local Similarity 51.28; Pred. No. 5e-33;
 Matches 84; Conservative 33; Mismatches 41; Indels 6; Gaps 3;

QY 7 TSSOKALML--ELKSIQEPVFCFRITLVDSILYNNHVAIFGLPNLYEGSYFKAHIK 64
 DB 4 TTSQASLLQKQLKALNKNVFGSAGLVDSDNVFEWCVTIIGPFTLYGGYNA:MST 63
 QY 65 PIDYPSPTFRFLTKMHIPNIYENGVCISILHPPVDDFGSGLFSPERNPTQNVRTI 124
 DB 64 PONTNSPPTFRFTSMHHPNYPDGRVCISILHPPGDDPNQYELASERWLPVHIVE: 123
 QY 125 LLSVISLNEPNTFSANVDASVMEKRWDSKGDKEYAEIIRK 167
 DB 124 LLSVISLNGDNDESANIEA---KEWRKRDHPKKRVPLVRK 164

RESULT 9
 Q42540 PRELIMINARY: PRI: 166 AA.

AC Q42540;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19) (Ubiquitin-protein
 DE ligase) (Ubiquitin carrier protein) (Ubiquitin-conjugating enzyme
 DE UBC7).
 GN UBC1 OR UBC7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SPECIFIC;
 RA MEDLINE=96218124; PubMed=8647407;
 RA van Nocker S., Walker J.M., Vierstra R.D.;
 RT "The Arabidopsis thaliana UBC7/13/14 genes encode a family of
 RT multiubiquitin chain forming E2 enzymes.";
 RL J. Biol. Chem. 271:12150-12158(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=99156233; PubMed=10048488;
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and YAC clones.";

Db	23	ASQSLLLQKLDLCKHPVGFSGAGLVDEKNITFWSVIFGGPDLYEGSGFENAIPIFFP	R2
Qy	66	IDYVSPPTREFLIKMMHPIYENGDCVCSILHPVDVDPQSGELPSEKRNPTQNVKILL	135
Db	83	QNVPSPTVKTSDMMHPIYSDGRVCISLIPQDHFISYELASDPTPVHVSFIML	142
Qy	126	SVISLLNPNPISPNVDASVMFKRWDGSKG-KREYAEILRK	167
Db	143	SIISMLSPNDSPANVEAA---KEWRDKRDEFKKVRKVRK	162
RESULT 1:			
ID	022656	PRELIMINARY:	PRF: 169 AA.
AC	022656		
DT	01-JAN-1998 (TrEMBLrel_05, Created)		
DI	01-JAN-1998 (TrEMBLrel_05, last sequence update)		
DE	01-MAR-2002 (TrEMBLrel_26, last annotation update)		
DE	Ubiquitin-conjugating enzyme protein E2.		
GN	UBC7.		
OS	Zea mays (Maize).		
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACF clade;		
OC	Panicoidae; Andropogonaceae; Zea.		
NCBI	NCBI_TaxID=4577;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN:ZE:		
RA	van Nocker S., Androva M., Schroeder J.L., Vierstra R.D.;		
RT	z-UBIC7-1, a ubiquitin-conjugating enzyme from maize expressed in		
RT	mitotically active tissue (Accession: No. AF032458) (PCR86-333);		
PL	Plant. Physiol. 116:1191-1191(1998).		
CC	!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER		
CC	PROTEINS (BY SIMILARITY).		
CC	!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +		
CC	DIPHOSPHATE + PROTEIN-N-UBIQUITYLYSINE.		
CC	!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.		
CC	!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-		
CC	THIOLESTER FORMATION (BY SIMILARITY).		
CC	!- SIMILARITY: BELONGS TO THE UBIQUITIN CONJUGATING ENZYME FAMILY.		
DR	EMBL: AF032458; AAC12662.1; -		
DR	HSSP: Q02159; 20C.		
DR	InterPro: IPR000608; JRG_ccs3gat.		
DR	Pfam: PF06179; UQ_ccc; 1.		
DR	PRODOM: PD000461; UBU_ccn3gat; 1.		
DR	SMART: SM00212; UBCG; 1.		
DR	PROSITE: PS0183; UBIQUITIN_CONJUGAL_1; 1.		
DR	PROSITE: PS0187; UBIQUITIN_CONJUGAL_2; 1.		
DR	Ligase; Ubiquitin conjugation.		
SK	SEQUENCE 169 AA; 18970 MW; CA46HDCB27584954 CRC64;		
Query Match 29.4% Score 444; DB 10; Length 169;			
Best local similarity 49.4%; Pred. No. 2.7e-32;			
Matches 81; Conservative 37; Mismatches 40; Indels 6; Gaps			
Qy	7	TSSQKALML--ELKSLGEEPEVGEFRITLVDSILYKFWVAIGLPIYEGYFKARKE	64
Db	4	SASQSLLLQKLDLCKHPVGFSGAGLVDEKNITFWSVIFGGPDLYEGSGFENAIPIFFP	63
Qy	65	PIDYVSPPTREFLIKMMHPIYENGDCVCSILHPVDVDPQSGELPSEKRNPTQNVKILL	124
Db	64	PQNVPSPTVKTSDMMHPIYSDGRVCISILHPQDHFISYELASDPTPVHVSFIML	142
Qy	125	LVISLLNPNPISPNVDASVMFKRWDGSKG-KREYAEILRK	167
Db	124	LSTISMLSPNDSPANVEAA---KWRDKRDEFKKVRKVRK	164
RESULT 2:			
ID	024560	PRELIMINARY:	PRF: 169 AA.
ID	024560		
ID	024560		

124 LSTLMSLSPNESPAN.FAA---KJWREKDESKKKVQCVPK 164

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Ubiquitin carrier protein.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-INBRED LINE A188;
 RX MEDLINE=99105148; PubMed=9890811;
 RA De Vries A., Cordts S., Dresselhaus T.;
 RT Molecular Characterization of a cDNA Encoding an Ubiquitin Carrier
 RT Protein (UBC7) Isolated from Egg Cells of Maize (Accession No. AJ002959
 RT PG98-177).
 RL Plant. Physiol. 118:1101-1101(1998).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AJ002959; CAA05772.1; -.
 DR HSSP: Q02159; 2UCZ.
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF001179; UQ_con: 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBCc; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAL_1; 1.
 DR PROSITE: PS50127; UBIQUITIN_CONJUGAL_2; 1.
 KW Ligase; Ubiquitin conjugation.
 SQ SEQUENCE 169 AA: 19151 MW: A96816EF4576B69B CRC64;
 Query Match 29.4%; Score 444; DB 10; Length 169;
 Best Local Similarity 49.4%; Pred. No. 2.7e-32;
 Matches 81; Conservative 37; Mismatches 40; Indels 6; Gaps 3;
 QY 7 TSSOKALML--ELKSQLEPVEGFRITLVESOLYNWFAVGLPNTLYEGGYKAHKF 64
 DB 4 TTTQASLLQKLRDLAKHPVDSFAGIVDSNVFQVITIGPPDTLYDGGYENAKMSF 63
 QY 65 PUDYPSPTFRFLTKMHPNLYENGVCISLHPVDDPOSGELPSRWNPIONVETIL 124
 DB 64 PONTNPPSPVRETSMHPNPVYDGRVCISLHPGCDPNGYELASRWTPVITVFSIV 123
 QY 125 LSVISLLNEPTSPANVDASVHKRWDKSGK-DKEYAEIRK 167
 DB 124 LSIISMSISNDSPANIEAA---KHWREQREDKKKVRIVRK 164
 RESULT 13
 Q42541
 ID Q42541 PRELIMINARY; PRT; 166 AA.
 AC Q42541;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19) (Ubiquitin-protein
 DE ligase) (Ubiquitin carrier protein) (AT3946460/v18i15_180).
 GN UBC13 OR F18i15_180.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE ETIOLATED; AND ETHYLENE TREATED SEEDLINGS;
 RX MEDLINE=96218124; PubMed=8647807;

RA van Nocker S., Walker J.M., Vierstra R.D.;
 RI "The Arabidopsis thaliana UBC7/13/14 genes encode a family of
 RI multi-ubiquitin chain-forming E2 enzymes";
 RI J. Biol. Chem. 271:12150-12158(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chotisme N., Robert C., Brothier P., Wincker P., Cattolico L.,
 RA Artiguenave F., Saurin V., Weissbach J., Mewes H.W., Lemcke K.,
 RA Mayor K.F.X., Quettier F., Salanoubat M.;
 RI Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RI Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.-C., Koesoma E., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Poir C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Sotowick A.,
 RA Tang C.-C., Tortum H., Yamada K., Yamamura Y., Yu C., Yu S.,
 RA Shinozaki K., Davis R.W., Tackoligis A., Ecker J.R.;
 RI "Arabidopsis cDNA clones";
 RI Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.-C., Meyers M.C., Shinn P., Banh J.,
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Tang C.-C., Tortum H., Wu H.-C., Yamada K.,
 RA Yamamura Y., Yu C., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RI "Arabidopsis cDNA clones";
 RI Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: Q33758; AAC49322.1; -.
 DR EMBL: AL133298; CAB62037.1; -.
 DR EMBL: AY050368; AAK9385.1; -.
 DR EMBL: AY094040; AAM16196.1; -.
 DR HSSP: Q02159; 2UCZ.
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF001179; UQ_con: 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBCc; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAL_1; 1.
 DR PROSITE: PS50127; UBIQUITIN_CONJUGAL_2; 1.
 KW Ligase; ubiquitin conjugation.
 FI BINDING 89 89 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 166 AA: 18822 MW: DCC26424275F275B CRC64;
 Query Match 29.1%; Score 439; DB 10; Length 166;
 Best Local Similarity 50.9%; Pred. No. 7.3e-32;
 Matches 83; Conservative 32; Mismatches 43; Indels 6; Gaps 3;
 QY 8 SSOKALML--ELKSQLEPVEGFRITLVESOLYNWFAVGLPNTLYEGGYKAHKF 65
 DB 2 NSQACLLQLQKLRDLAKHPVDSFAGIVDSNVFQVITIGPPDTLYDGGYENAKMSF 61
 QY 66 IDYPSPTFRFLTKMHPNLYENGVCISLHPVDDPOSGELPSRWNPIONVETIL 125
 DB 62 QNPNPPTVRFRTSD:WHNPVDFGVGVCISLHPGCDPNGYELASRWTPVITVFSIV 121


```

DB 64 DYPLMPKMAFTTEIMHNPVNGEVCISILHPDGDYKYEINAGERWLPVHSPETILIS 123
QY 127 VISLNEPTFSPANDVMEKRWOSKGDKEYAEILRKQV 169
DB 124 VISMSLSPNDSPANDAAKERE-----NPQPKRVRRIIV 160
NCBI_TaxID=5141;

RESULT 16
Q9C2A5 PRELIMINARY: PRT: 166 AA.
AC Q9C2A5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Probable ubiquitin-conjugating enzyme ubcF3.
GN BLIN2.230.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Hartmann B., Holland R.,
RA Nyakatura G., McWee H.W., Mannhaupt G.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDIJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDIJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: AL513444; CAC28704.1;
DR HSSP: Q02159; 2UCZ.
DR InterPro: IPR000608; UBO_conjugat.
DR Pfam: PF001179; UO_con; 1.
DR ProDom: PD000461; UBO_conjugat; 1.
DR SMART: SM00212; UBCF; 1.
DR PROSITE: PS0183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; ubiquitin conjugation.
SQ SEQUENCE 166 AA; 18540 MW; 2387742CA5477185 CRC64;

Query Match: 26.38; Score 398; DB 3; Length 166;
Best Local Similarity 48.88; Pred. No. 3.7e-28;
Matches 80; Conservative 25; Mismatches 53; Indels 6; Gaps 2;

QY 9 SQAALMLKLSKLOFPPVPCFRTLLVDSEDLNKNFVAIFGLPNILYEGGYFAHKKFPNDY 68
DB 6 AQRLEQFRAITNNPPGIFTAGTSPISDILLRWELIOGPECTPFGGVFAEILKPNFY 65
QY 69 PYSPTPFLTKMHPNIYENGVCISILHPVDQDQSGELPSEWNPNTQVRIILSVI 128
DB 66 PHMPTMKFLGDTFHPNYPSCISILHPDGDNDPHYETASERWSTQSVKILLISVM 125
QY 129 SILNEPNTSPANDVMEKRWOSKGDKEYAEILRKQVSAI 172
DB 126 SHLAEPNDSPANVEAKH---WRE---KRSEYEARVKASVRAS 163

RESULT 17
Q9VXE8 PRELIMINARY: PRT: 167 AA.
AC Q9VXE8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

```

```

DE CG4443 protein (RE34144p).
GN CRU OR CG4443.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=16731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans G.A., Gocayne J.D.,
RA Aramantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazej R.G., Champe M., Pfeiffer R.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.S.,
RA Haller K.M., Basu A., Baxendale J., Haytrakaroglu I., Beasley E.M.,
RA Becson K.Y., Horos P.V., Berman B.P., Bhandari B., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Bartis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chanira T.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dudan-Rochia S., Dunkov H.C., Egan P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garq N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibequm C.,
RA Jaimel B., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimball M.R., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson B.,
RA Merkulov G., Milstine N.V., Mobarry C., Morris J., Mosureff A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzey D.M., Nelson B.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.M., Pacich J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.S.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong K., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang S., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter G.C.;
RT The genome sequence of Drosophila melanogaster.
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chaver C., Dorsett V., Dresnek D., Farfan D., Friso E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nanoo J., Pacich J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celisner S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDIJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: AE003502; AAF48626.1;
DR EMBL: AY071319; AAL48941.1;
DR HSSP: Q02159; 2UCZ.
DR FlyBase: FBgn0015374; crl.
DR InterPro: IPR000608; UBO_conjugat.
DR Pfam: PF00179; UO_con; 1.

```

Q9N925	PRELIMINARY	PRT: 200 AA.
ID	Q9N925	
AC	Q9N925;	
DT	01-OCT-2000 (TREMBLrel.15, Created)	
ST	01-OCT-2000 (TREMBLrel.15, last sequence update)	
ST	01-MAR-2002 (TREMBLrel.20, last annotation update)	
DE	Ubiquitin conjugating enzyme.	
GN	UBL5 OR UBC7 OR CG4443.	
OS	Drosophila melanogaster (fruit fly).	
CS	Drosophila: Melazoza; Arthropoda; Insecta; Hexapoda; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI_TaxID=7227;	
RI	[1]	
RN	SEQUENCE FROM N.A.	
RP	MEDLINE=20341328; PubMed=10880487;	
RA	Orquad S., Rosenfeld G., Greenspan K.G., Seegal D.;	
RX	*courtless, the drosophila UBC7 homolog, is involved in Male Courtship	
RI	Behavior and Spermatogenesis.*;	
RJ	Genetics 155:1267-1280(2000).	
RT	!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER	
CC	PROTEINS (BY SIMILARITY).	
CC	!- CATALYTIC ACTIVITY: ATP - UBIQUITIN -> PROTEIN LYSINE + AMP +	
CC	DIPHOSPHATE -> PROTEIN N-UBIQUITYLYSINE.	
CC	!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.	
CC	!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-	
CC	THIO-ESTER FORMATION (BY SIMILARITY).	
CC	!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.	
CC	PMBL: AJ277746; CAB90824.1; ..	
DR	HSPP; Q02159; 2UCZ;	
DR	Flybase; Fban0315374; cbl.	
DR	InterPro; IPR000608; Cblq_conjugat.	
DR	Pfam; PF03179; UQ_con; 1.	
DR	ProDom; PD000461; UHQ_conjugat; 1.	
DR	SMART; SM00212; UBC; 1.	
DR	PROSITE; PS00183; UBIQUITIN_CONJUGAL_1; 1.	
DR	PROSITE; PS0127; UBIQUITIN_CONJUGAT_2; 1.	
DR	Ligase; ubiquitin conjugation.	
KW	SEQUENCE 200 AA; 23377 MW; 9041932AD5C204B CIRC04;	

```

Query Match      25.3%, Score 382, DR 5, Length 200;
Best Local Similarity 50.7%, Pred. No. 1.3e-26;
Matches 74; Conservative 21; Mismatches 51; Indels 0; Gaps 0;

QY      8 SSOALMIEKLSLOEPPVGFHNIIVDSQLYNWEVAIKLLNTLIYGGVEKAHKKYID 67
DB      4 SAKRLRLMAEYKQITLDPHGIVAGDISDMNFFWEALIASPNTCFSGGVIPAKLIFPFS 63

QY      68 VYSPPTRELTAKMHIPNIYGVDCISILHPVDVDPQSGEIPSEKNPTQNTVPTILSV 127
DB      61 MAY-2000 (TREMBLrel. 13, last sequence update)
DB      64 VYSPPKMKTCOMEFHPIFAIKRVCICTIHAGDQPMGVSLSAKSEFVQSVKILLSV 123

QY      128 ISLLNPENTSPANVDASVMEKHWKD 153
DB      224 VSLAEZPDSGANVDAAIMKREQD 149

RESULT 20
Q9U1Q1
AC Q9U1Q1 PRELIMINARY: PRI; 170 AA.
ID Q9U1Q1
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DI 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DI 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Y87G2A.9 protein.
GN Y87G2A.9
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; rhabditinae; Caenorhabditis.
OX NCBI_TaxID=5239;
[2]
RN RP SEQUENCE FROM N.A.
RA White S.;

```

Q90101	PRELIMINARY:	PRT:	170 AA.
ID	Q90101		
AC	Q90101		
DT	01-MAY-2000	(TRENBLER)	13, Created
DI	01-MAY-2000	(TRENBLER)	13, Last sequence update
DE	01-MAR-2002	(TRENBLER)	20, Last annotation update
DE	Y87GA.9	protein.	
GN	Y87GA.9		
OS	Caeoarthabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhadida; Oligo-		
OC	Rhabditidae; Peloderiinae; Caeoarthabditis.		
OX	NCBI_TaxID=6239;		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	White S.		

Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RL SEQUENCE FROM N.A.
 RP MEDLINE-990696113; PubMed-9851916;
 RA none;
 RX none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology.";
 RL Science 282:2012-2018(1998).
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: AL110500; CAB60431.1; .
 DR HSSP: Q02159; 20CZ.
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00179; UBQ_con; 1.
 DR ProDom: PD000461; UBQ_conjugat; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
 KW ligase; ubiquitin conjugation.
 SQ SEQUENCE 170 AA; 19056 MW; 0A4C0RE49D8A4B4D CRC64;

Query Match 25.2%; Score 381; DB 5; Length 170;
 Best Local Similarity 46.4%; Pred. No. 1.3e-26;
 Matches 77; Conservative 24; Mismatches 61; Indels 4; Gaps 2;

QY 11 KALMLELKSQEPVRCFRITLVDSGLYNWEVAIFGLPNTLYEGYKAHKEIDYVPY 70
 Db 7 KRLMEYKELTIRPEGLIAAPIDEDNFENFEWELTGPEETCFANVFPARTFPDDIPL 66
 QY 71 SPTPTFRITKMHNIYENGDCVCSILHPPVDDPQSGELPSRWNPNTONVRITLLSVTSI 130
 Db 67 SPKMRHFCGIFHNVAIDCRVCISILHAPGDDPTGVLSNERNVSPQSIKILLSVSM 126
 QY 131 LNEPNTSPANVDASYMRKWRDCKOKKAE-ITRKQVSATKAE 175
 Db 127 LAEPNDESPANVSAAM---WRDRAQPEKTADSLVRKTLCLPASE 169

RESULT 21
 Q8X0N3
 ID Q8X0N3 PRELIMINARY; PRT: 257 AA.
 DT 01-MAR-2002 (TEMBLrel. 20, Created)
 DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
 DE Probable ubiquitin-conjugating enzyme CDC34.
 GN B10H4.060.
 OS Neurospora crassa.
 CC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;
 CC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann H., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL670010; CAB21393.1; .
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00179; UBQ_con; 1.
 DR ProDom: PD000461; UBQ_conjugat; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
 SQ SEQUENCE 257 AA; 29134 MW; CF53F44A7A9BC295 CRC64;

Query Match 24.7%; Score 373.5; DB 3; Length 257;
 Best Local Similarity 36.3%; Pred. No. 1.1e-25;
 Matches 93; Conservative 41; Mismatches 75; Indels 49; Gaps 10;

QY 13 IMLELKSQEPVRCFRITLVDSGLYNWEVAIFGLPNTLYEGYKAHKEIDYVPY 49
 Db 14 IMSEKALKEKNTHTFHVRVRSIVRFTSFAGLSFACADDSATLNIKWR---PELMVNP 79
 QY 50 NTLVEGYEKAHKEIDYVPYSPTEFRELTKMHNIYENGDCVCSILHPPVDDPQSGEL 109
 Db 71 DSAENGCPYKACWESDRYPQPKFHFLLIPTHPNVYDQICLSILHFGEDLMSGQ 130
 QY 110 PSRWNPNTONVRITLLSVTSIENPNTFSANVDASVYMRKWRDCKOKKAEITRKQV 169
 Db 131 ASERWSPLOGAESVLRSLVLLIDUPEINSPANVDAGVYRERRE-----LYNKKARETV 184
 QY 170 SAJKAERKDGKVPITLAEVCITKVP-----SNQN-----SSQLLYD---GA YDD:DEE 218
 Db 185 ERSKCT-PPGFEMVSPFE-----KPPKAENDDDFWAESESEFEGSGSDGDDDFMEL 236
 QY 219 DEEDADCYDDDDSGMR 235
 Db 239 FEGDGESEDDDDSDNSNK 256

RESULT 22
 Q9OX59
 ID Q9OX59 PRELIMINARY; PRT: 148 AA.
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE Ubc7p homolog.
 GN YBE2G2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lenk U., Sommer T., Hartmann E.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL: U93241; AAF2503.1; .
 DR HSSP: Q02159; 20CZ.
 DR MGD: MGI:1343188; Ube2g2
 DR InterPro: IPR000608; UBQ_conjugat.
 DR Pfam: PF00179; UBQ_con; 1.
 DR ProDom: PD000461; UBQ_conjugat; 1.
 DR SMART: SM00212; UBCC; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1;
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1;
 KW ligase; Ubiquitin conjugation.
 SQ SEQUENCE 148 AA; 16677 MW; C544R92AD7CA3857 CRC64;

Query Match 23.8%; Score 359.5; DB 11; Length 148;
 Best Local Similarity 53.7%; Pred. No. 9.4e-49;
 Matches 74; Conservative 18; Mismatches 47; Indels 7; Gaps 2;

QY 24 PVGEPRITLVDSGLYNWEVAIFGLPNTLYEGYKAHKEIDYVPYSPTEFRELTKMHKE 83
 Db 3 PPEGIVAGPMNEENFEWELNGLMGPEDTCTFEVGFPAISLEPLDPLSPKMRFTCEMFH 62
 QY 84 PNLYENGDCVCSILHPPVDDPQSGELPSRWNPNTONVRITLLSVTSIENPNTFSANVD 143


```

Db 63 PNYPDGRVCSILHARGDPMGYESSAERWSPQSVKEKLLSVSMCAEPNDESGANDV 122
QY 144 ASVMPFKNRSGKGRKVAHRIKQV 169
Db 123 ASKM---NRD-----DREQFYKIAKQI 141

RESULT 23
Q9CZJ4 PRELIMINARY: PRT: 136 AA.
AC Q9CZJ4
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 2700059C12Hik protein.
GN 2700059C12HIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1.
RX MEDLINE=21085660; PubMed=11217851;
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojibori T., Hono H., Kasukawa T., Saito K.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner I., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Raldarelli R., Barsh G.,
RA Blake J., Roffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein H.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez J., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Joyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
RA Hayashizaki Y.;
RI Functional annotation of a full-length mouse cDNA collection.*;
HL Nature 409:685-690(2001).
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.
CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: AK012539; BAB28305.1; -.
DR HSP: Q02159; 20C2.
DR MGD: MGI:1914378; 2700059C12RIK.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UBQ_conj_1.
DR Prodom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBCC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 136 AA; 15605 MW; D50282F47F4A2H9F CRC64;

```

Query Match 21.0%; Score 317.5; DB 11; Length 136;
 Best Local Similarity 49.2%; Pred. No. 5.2e-21;
 Matches 58; Conservative 23; Mismatches 32; Indels 5; Gaps 2;

```

QY 52 IYEGCYFAHKKIPDIPYSPPTFELTKMHPNINYGVDCISILHPPVDPOSSELP 111
Db 15 IHFGCVFAHLLTFPKOYPLRPKPKMEITEIHPNPNVDKNGDVCISILHGEDKGYEKPE 74

```

```

QY 112 FRNPNTCNVRTILSLVSLNEPNTFSPANVDASVNFKRWKSKKC--KEYAEIIRK 167
Db 75 ERWLP-HTVETLIMISVLSADPNDSNPANVDAA--KFWHRDNGEKKKVCYRK 129

RESULT 24
Q8SSK3 PRELIMINARY: PRT: 216 AA.
AC Q8SSK3
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ubiquitin conjugating enzyme E2.
GN EC0011010.
OS Encephalitozoon cuniculi.
OC Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574510; PubMed=11719836;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Ihomarai F.,
RA Prensier G., Harbe V., Peyrotailade E., Brotier P., Wincker B.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gossy M.,
RA Weissenbach J., Vivares C.P.;
RT Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.*;
HL Nature 414:450-453(2001).
DR EMBL: AL391737; CA024572.1; -.
SQ SEQUENCE 216 AA; 25055 MW; 9H45EFD6C17C0F0 CRC64;

Query Match 20.2%; Score 304.5; DB 5; Length 216;
Best Local Similarity 38.5%; Pred. No. 1.5e-19;
Matches 69; Conservative 30; Mismatches 51; Indels 29; Gaps 6;

QY 17 IKSLOEPVEGFRITLVDSLYN-----WEVAIFGLPNILYEGYFKAF 62
Db 31 LRLESNEE-ER-KIADGDKLFSAFYPRGSMENRDKYWDIYFTLGSDSLAGRLKAVM 59
QY 63 KPPTDIPYSPPTFELTKMHPNINYGVDCISILHPPVDPOSSELP--PSRWKNTON 119
Db 90 KFPSSYPLRPPTLKFVSKMFEFNLVECKMQLSLEF-DKQSSSVFQSPKDKWIPVCN 147
QY 120 VRTILSLVSLNEPNTFSPANVDASVNFKRWKSKGDKHYAEINKYVSKATKAEK 176
Db 148 IRTIVMSIIVLISNPISNPANVDASVNYR-----UNPERY---KEVRIARPEREK 197

RESULT 25
Q8SR17 PRELIMINARY: PRT: 162 AA.
AC Q8SR17
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ubiquitin conjugating enzyme E2-17kDa.
GN EC01013101.
OS Encephalitozoon cuniculi.
OC Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;

```

DE	FMBL:	AKG11363;	BAB27579.2; 100%
DR	HSSP:	P29865;	ZAAK.
DR	MGD:	MG1:102944;	Ube2b.
DR	IcterPro:	LPRQ00608;	UBO_conjugat.
DR	pfam:	PF00179;	UQ_conj_1.
DR	ProDom:	PD000461;	UBO_conjugat; 1.
DR	SMART:	SM00212;	UBCC; 1.
DR	PROSITE:	PS00183;	UBIQUITIN_CONJUGAT_1; 1.
DR	PROSITE:	PS50127;	UBIQUITIN_CONJUGAT_2; 1.
KW	Ligase:	Ubiquitin conjugation.	
SC	SEQUENCE	152 AA:	17285 MW, 3EB728C853B40A6 CRC64;
 Query Match Best Local Similarity 19.5%; Score 294; DB 1; Length 152; Matches 65; Conservative 31; Mismatches 53; Indels 24; Gaps			
OY	7 TSSOKA:	MLELASLOEPVGEGR-TLVDSLDYNKEVA;FGLNTLYEGCYFAH;KFPI 66	
DB	3 TPARRRIMDFKRLQEDPPAGVS-GAPSENNIMVNNAVIFGPSTIGDTFKLVTFSE 61		
OY	67 DYPSPTFRFTTKMHNIENGVDCISILHPVDHQSDCLPSRWKPDPONVTILLS 126		
DB	62 EYPNKPPTVEFLSKPHFNVAQSICDILL-----QARNSPYDVSSLLS 168		
OY	127 VISLNLPNTFSPANVDASVMFERKWRSKGKKEYAEIRIKQVSA;KAERKD 179		
DB	109 IQSLIDPEPNIPANSQAOLYQR-----NKREV---HKRVSAIVGSND 151		
 RESULT 27 Q96FX4 PRELMINARY; PRT; 152 AA. ID Q96FX4 AC Q96FX4: CD 01-DEC-2001 (IREBLrei.19, Created) DI 01-DEC-2001 (IREHLrei.19, Last sequence update) DE 01-MAR-2002 (IREBLrei.20, Last annotation update) TE Ubiquitin-conjugating enzyme E2A (RAD6 homolog). OS Homo sapiens (Human). OC Fukuyama; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Butterfaia; Primates; Catarrhini; Hominiidae; Homo. NCBI_TaxId=9606; RN [1] RX SEQUENCE FROM N.A. RC TRUSS-SKIN, RP Strausberg R.; RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases. CC -! FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS (BY SIMILARITY). CC -! CATALYTIC ACTIVITY: APC + UBIQUITIN + PROTEIN LYSINE + AMP = DIPHOSPHATE + PROTEIN N-UBIQUITYL-LYSINE. CC -! PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION. CC -! MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOESTER FORMATION (BY SIMILARITY). CC -! SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY. CC EMBL: BC010175; AAAH0175.1; DR InterPro: IPR006038; UBQ_conjugat. DR pfam: PF00179; UQ_conj_1. DR ProDom: PD000461; UBQ_conjugat; 1. DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; UNKNOWN_1. DR PROSITE: PS50127; UBIQUITIN_CONJUGAT_2; 1. KW Ligase: Ubiquitin conjugation. SC SEQUENCE 152 AA: 17315 MW, 0AAE59R7770E47E2 CRC64;			
 Query Match Best Local Similarity 19.4%; Score 293; DB 4; Length 152; Matches 64; Conservative 33; Mismatches 52; Indels 24; Gaps			
OY	7 TSSOKA:	MLELASLOEPVGEGRITLVDSNLKNWEVA;FGTLTYEGCYFAH;KFPI 66	
DB	3 TPARRRIMDFKRLQEDPPAGVS-GAPSENNIMVNNAVIFGPSTIGDTFKLVTFSE 61		
OY	67 DYPSPTFRFTTKMHNIENGVDCISILHPVDHQSDCLPSRWKPDPONVTILLS 126		

```

Db 62 EYPNKPPTVRFVSKMFUPNYVYAGSICLDIL-----QNRWSPYDVSSILTS 108
QY 127 VISLLNEPNTFSPANVDASVMFKWRDSCGKDKKEYAEIIRKQVSATKAEAKD 179
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 IOSLLDEPNPNSPANQAOLYQE-----NKREY-----EKRVSAIVEQSWRD 151

RESULT 28
Q92255 ID Q92255 PRELIMINARY; PRT; 152 AA.
AC Q92255;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ubiquitin-conjugating enzyme HR6A (Ubiquitin-conjugating enzyme E2A,
GN RAD6 homolog) (S. cerevisiae).
DB UBE2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BCBA; TISSUE=BRAIN;
RA Roest H.P., van Klaveren J., Koken M.H.M., Vermeij M.,
RA van Cappellen W.A., Baarends W.M., Hoogerbrugge J.W., Bootsma D.,
RA Hoelmakers J.H.J., Grootegeed J.A., de Wit J.;
RT "Isolation of mHR6A, a gene highly homologous to the male fertility
RT gene mHR6B.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=SKELETAL MUSCLE;
RA Kwon Y.T., Varshavsky A.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: AF089812; AAC64563.1;
DR EMBL: AF383148; AAK62984.1;
DR EMBL: BC026053; AAB26053.1;
DR HSP: P25865; ZAAK.
DR MGD: MGI:102959; Ube2a.
DR InterPro: IPR000608; URQ_conjugat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; URQ_conjugat; 1.
DR SMART: SM00212; UBCc; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 152 AA; 17315 MW; 0AAEB5B7770E47E2 CRC64;

Query Match 19.4%; Score 293; DB 11; Length 152;
Best Local Similarity 37.0%; Pred. No. 9.8e 19;
Matches 64; Conservative 33; Mismatches 52; Indels 24; Gaps 4;

QY 7 TSSQKALMLKLSLOEPEVEGRITLVDSLYNWEVAIFGLPNTLYEGGYFKAHIKFPI 66
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 TPARRRLMRDEKRLQEDPPAGVS-GAPSENNIMYNAVIFGPGEGITFKLTIEFT 61

QY 67 DYPYSPPTFRFLKMMHHPNLYENGIVGICISILHPVDPDSGLPSEFMWNPQNTVITLLS 126
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 EYPNKPPTVRFVSKMFUPNYVYAGSICLDIL-----QNRWSPYDVSSILTS 108

QY 127 VISLLNEPNTFSPANVDASVMFKWRDSCGKDKKEYAEIIRKQVSATKAEAKD 179
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 IOSLLDEPNPNSPANQAOLYQE-----NKREY-----EKRVSAIVEQSWRD 151

RESULT 30
Q960P5 ID Q960P5 PRELIMINARY; PRT; 151 AA.
AC Q960P5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE UVS2.
GN UVS2.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]

```

```

QY 127 VISLLNEPNTFSPANVDASVMFKWRDSCGKDKKEYAEIIRKQVSATKAEAKD 179
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 IOSLLDEPNPNSPANQAOLYQE-----NKREY-----EKRVSAIVEQSWRD 151

RESULT 29
Q9W6F3 ID Q9W6F3 PRELIMINARY; PRT; 152 AA.
AC Q9W6F3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Ubiquitin-conjugating enzyme.
DE Ubiquitin-conjugating enzyme.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sofer L., Butnside J.;
RT "Chicken ubiquitin-conjugating enzyme.";
RL Submitted (JAN-1999) to the EMBL/GenBank/CCRC databases.
CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR EMBL: AF120212; AAO31646.1;
DR HSP: P25865; ZAAK.
DR InterPro: IPR000608; URQ_conjugat.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; URQ_conjugat; 1.
DR SMART: SM00212; UBCc; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 152 AA; 17315 MW; 0AAEB5B7770E47E2 CRC64;

Query Match 19.4%; Score 293; DB 13; Length 152;
Best Local Similarity 37.0%; Pred. No. 9.8e 19;
Matches 64; Conservative 33; Mismatches 52; Indels 24; Gaps 4;

QY 7 TSSQKALMLKLSLOEPEVEGRITLVDSLYNWEVAIFGLPNTLYEGGYFKAHIKFPI 66
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 TPARRRLMRDEKRLQEDPPAGVS-GAPSENNIMYNAVIFGPGEGITFKLTIEFT 61

QY 67 DYPYSPPTFRFLKMMHHPNLYENGIVGICISILHPVDPDSGLPSEFMWNPQNTVITLLS 126
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 EYPNKPPTVRFVSKMFUPNYVYAGSICLDIL-----QNRWSPYDVSSILTS 108

QY 127 VISLLNEPNTFSPANVDASVMFKWRDSCGKDKKEYAEIIRKQVSATKAEAKD 179
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 IOSLLDEPNPNSPANQAOLYQE-----NKREY-----EKRVSAIVEQSWRD 151

RESULT 30
Q960P5 ID Q960P5 PRELIMINARY; PRT; 151 AA.
AC Q960P5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE UVS2.
GN UVS2.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiates; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]

```

DR FMRL: U06521; AAB47850.1; .
DR HSSP: P06104; .AYZ.
DR InterPro: IPR000608; UBQ_conjugat.
DR Pfam: PF00179; UQ_con: 1
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBQC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1;
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1;
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 151 AA; 17200 MW; 157024N4EA568CCF CRC64;

Query Match 18.7%; Score 283; DB 3; length 151;
Best Local Similarity 35.0%; Pred. No. 7.8e-18;
Matches 57; Conservative 36; Mismatches 50; Indels 20; Gaps

QY 7 ISSSKALMLKLSLOEPEVEGFRITIVDESDIYNNEVAIVGLNTILYEGSYFKAHIKETPI 56
DB 3 TAARRLRMRKFKKMTDPPAGVSASPPVC-NVMIINAVTGGPATITFENGTFRLVMGEE 63
QY 57 DYPXSPPTFRDLKKHNNIYENGDCVTSIIHPVDQPSGELPSRWNPIONKRIILIS 126
DB 52 QYNNKPQVQKFISEHFNNAVATGELDLIL-----QNRKSPFYDVAAVETS 108
QY 127 VISLNLKPNTFSPANDVASVFRKWRKSKDKKEYAKHIRKQV 169
DB 109 TOSLNDPNTGSPANVEASNIY-----KDRKKEYTKKRVET 145

RESULT 32
Q9AVP0 PRELIMINARY; PRI: 152 AA.
AC Q9AVP0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Ubiquitin-conjugating enzyme (E2).
GN NRUBCL.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
OC Asteridae; eustersids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxId:4097;
RI [1]
RN SEQUENCE FROM N.A.
RC STRAIN-CV. SAMSUN 11;
RA Koyama T., Okada T., Kitajima S., Takagi M., Saishi H., Sato F.;
RT *Isolation of protein(s) which interact with ERF3 in cultured tobacco
RI cells**;
RI Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC PROTEINS (BY SIMILARITY).
CC !- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE -> AMP +
CC DIPHOSPHATE + PROTEIN-N-UBIQUITYLSINE.
CC !- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC !- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC THIOLESTER FORMATION (BY SIMILARITY).
CC !- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC EMBL: AB026055; BAB40310.1; .
DR HSSP: P25865; 2NAB.
DR InterPro: IPR000608; UBQ_conjugat.
DR InterPro: IPR001580; W34C.
DR Pfam: PF00179; UQ_con; 1.
DR ProDom: PD000461; UBQ_conjugat; 1.
DR SMART: SM00212; UBQC; 1.
DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1;
DR PROSITE: PS0127; UBIQUITIN_CONJUGAT_2; 1;
DR PROSITE: PS0678; WD_REPEATS_1; UNKNOWN_1;
KW Ligase; Ubiquitin conjugation.
SQ SEQUENCE 152 AA; 17355 MW; C02B66228C02032; CRC64;

Query Match 18.5%; Score 280; DB 30; length 152;
Best Local Similarity 38.0%; Pred. No. 1.5e-17;
Matches 62; Conservative 31; Mismatches 50; Indels 25; Gaps

QY 7 TSSOKALMLELKSLOEPVEGFRITLVDESGLYNWEVAIFGLPNTLYEGGYFKAHKFP1 66
 DB 3 TPARKRLMDFKRLQODPPAGTSGAPYD-NNIMLNAVIEGDDTPMDGGTFKLTLOFSE 61
 QY 67 DYPSPPTFRFLTKMHPNIVYENGVCISILHPVDDPQSGELPSERNWPTQNVRTILLS 126
 DB 62 DYPNKPPTFRFLTKMHPNIVYENGVCISILHPVDDPQSGELPSERNWPTQNVRTILLS 108
 QY 127 VISLNEPNTFSPANVDASVMFKWRDCKGKKEYAEIRKOV 169
 DB 109 IQSLCDPNPNSPANSEARME-----SENK-REYNRRVREIV 145
 RESULT 33
 Q9M4R0 PRELIMINARY; PRT; 152 AA.
 AC Q9M4R0
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Ubiquitin-conjugating protein.
 GN UBC.
 OS Avicennia marina.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Avicenniaceae; Avicennia.
 OX NCBI_TaxID=82927;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Parani M., Parida A.
 RT *Cloning of a cDNA coding for ubiquitin binding enzyme from the
 RT mangrove species Avicennia marina.*
 ML Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC EMBL: AF262934; AAF73016.1;
 CC HSSP: P25865; 2AAK.
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBC; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
 KW Ligase; Ubiquitin conjugation.
 SQ SEQUENCE 152 AA: 17362 MW; 7F8E8246C5310449 CRC64;

Query Match 18.3%; Score 277; DB 10; Length 152;
 Best Local Similarity 36.8%; Pred. No. 2,7e-17;
 Matches 60; Conservative 33; Mismatches 50; Indels 20; Gaps 4;

QY 7 TSSOKALMLELKSLOEPVEGFRITLVDESGLYNWEVAIFGLPNTLYEGGYFKAHKFP1 66
 DB 3 TPARKRLMDFKRLQODPPAGTSGAPYD-NNIMLNAVIEGDDTPMDGGTFKLTLOFSE 61
 QY 67 DYPSPPTFRFLTKMHPNIVYENGVCISILHPVDDPQSGELPSERNWPTQNVRTILLS 126
 DB 62 DYPNKPPTFRFLTKMHPNIVYENGVCISILHPVDDPQSGELPSERNWPTQNVRTILLS 108
 QY 127 VISLNEPNTFSPANVDASVMFKWRDCKGKKEYAEIRKOV 169
 DB 109 IQSLCDPNPNSPANSEARME-----SENK-REYNRRVREIV 145
 RESULT 34

Q9AVN9 PRELIMINARY; PRT; 152 AA.
 AC Q9AVN9
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Ubiquitin-conjugating enzyme (E2).
 GN NTUBC2.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koyama T., Kitajima S., Takagi M., Shishi H., Sato F.
 RT *Isolation of protein(s) which interact with EMB3 in cultured tobacco
 RT cells.*
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
 CC -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 CC EMBL: AB026056; BAB49311.1;
 CC HSSP: P25865; 2AAK.
 DR InterPro: IPR000608; UBC_conjugat.
 DR Pfam: PF00179; UQ_con; 1.
 DR ProDom: PD000461; UBC_conjugat; 1.
 DR SMART: SM00212; UBC; 1.
 DR PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
 KW Ligase; Ubiquitin conjugation.
 SQ SEQUENCE 152 AA: 17323 MW; CAZAG622933CF97 CRC64;
 Query Match 18.3%; Score 276; DB 10; Length 152;
 Best Local Similarity 37.4%; Pred. No. 3.4e-17;
 Matches 61; Conservative 32; Mismatches 50; Indels 20; Gaps 4;

QY 7 TSSOKALMLELKSLOEPVEGFRITLVDESGLYNWEVAIFGLPNTLYEGGYFKAHKFP1 66
 DB 3 TPARKRLMDFKRLQODPPAGTSGAPYD-NNIMLNAVIEGDDTPMDGGTFKLTLOFSE 61
 QY 67 DYPSPPTFRFLTKMHPNIVYENGVCISILHPVDDPQSGELPSERNWPTQNVRTILLS 126
 DB 62 DYPNKPPTFRFLTKMHPNIVYENGVCISILHPVDDPQSGELPSERNWPTQNVRTILLS 108
 QY 127 VISLNEPNTFSPANVDASVMFKWRDCKGKKEYAEIRKOV 169
 DB 109 IQSLCDPNPNSPANSEARME-----SENK-REYNRRVREIV 145
 RESULT 35
 Q95L52 PRELIMINARY; PRT; 124 AA.
 AC Q95L52
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE RAD6-like protein HR6A (Fragment).
 GN HR6A.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]

RA	Maylor P.J., Sisk F., Cawthra J., Handley P., Vogt G., Robertson L.,
RL	McDonagh P., Ivens A., Nguyen D., Munden H., Stuart K., Worthey E.A.
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBS databases.
CC	-!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC	PROTEINS (BY SIMILARITY).
CC	-!- CATALYTIC ACTIVITY: ATP + UBIQUITIN -> PROTEIN-LYSINE AMP +
CC	DIPHOSPHATE + PROTEIN-N-UBIQUITYLYNE.
CC	-!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC	-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC	THIOLESTER FORMATION (BY SIMILARITY).
CC	-!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR	EMBL: ALA45678; CAC14238.1; ..
DR	EMBL: AC022284; AAC10797.1; ..
DR	HSP: P15731; LOCQ.
DR	InterPro: IPR003608; UBQ_conjugat.
DR	Pfam: PF001179; UQ.con: 1.
DR	ProDom: PD000461; URO_conjugat; 1.
DR	SMART: SM00212; UBCG: 1.
DR	PROSITE: PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR	PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.
DR	Ligase: Ubiquitin conjugation.
KW	SEQUENCE 148 AA; 16651 MW; 797B2BA3C153B266 CRC64;
SQ	
Query Match	17.9%; Score 270; DR 5; Length 148;
Best Local Similarity	35.0%; Pred. No. 1.1e-16;
Matches 55; Conservative 29; Mismatches 55; Identities 18; Gaps	
QY	11 KALMLELKSLQEPPVERFRITLVDESDLYNNHVALFGPNTLVGGYFKAHKTPICYDPY 70
DB	4 RRIQELKULKOLEKPPANTIPSGPSVRSDELNKKATIIGHPHSYAAGLFELHHPPSYDF 63
QY	71 SPVPFRTLIKMMHPNIYENGDCVISILHPVDPPQSGLPSERWNPUNVTILLISVLS 130
DB	64 KPPLKLOFTIKIHPHNNGSGICLDIL-----KQNSPALTIISKVLSSVCSL ILQ
QY	131 LNEPNHFSPANVDASVHFKWRKSKK-----DREYA 162
DB	111 :TUNPDPIPVDTAKRYKTORNAFNKAAEWIHQYA 147
RESULT 37	
Q9NX15	
ID	Q9NX15 PRELIMINARY; PRT: 124 AA.
TC	Q9NX15;
DI	01-OCT-2000 (TREMBLrel. 15, Created)
DI	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DI	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE	RAD6 homolog (Fragment).
GN	BHR6A.
OS	Bos taurus (Rovine).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Bovinae; Bos.
OX	NCHI_TaxID=9913;
XP	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=OVARY.
CA	Pollumienko A.; Blecher S.;
RT	"H-X antigen as a tool for X and Y sperm separation.";
RL	Submitted (MAR-1999) to the EMBL/GenBank/DDA databases.
CC	-!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
CC	PROTEINS (BY SIMILARITY).
CC	-!- CATALYTIC ACTIVITY: ATP + UBIQUITIN -> PROTEIN-LYSINE AMP +
CC	DIPHOSPHATE + PROTEIN-N-UBIQUITYLYSINE.
CC	-!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC	-!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
CC	THIOLESTER FORMATION (BY SIMILARITY).
CC	-!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR	EMBL: AF132075; AAC36530.1; ..
DR	HSP: P25865; ZANK.
DR	InterPro: IPR000608; UBQ_conjugat.
DR	Pfam: PF001179; UQ.con: 1.
DR	ProDom: PD000461; URO_conjugat; 1.

QY	76	RELTKMHNENTYENDVCISILHPVDDPQSGELPSRKWNPNTQNRVAILLSVLSLNKP	13
DB	52	RFVSKMHPNVTADGSIQCLDI	1
QY	136	IFSPANDASVMFHK	150
DB	109	PNSPANSQAQVQYE	123
RESULT 39			
Q94AD2			
ID	Q94AD2	PRELIMINARY:	PRT: 145 AA.
AC	Q94AD2		
DT	01-DEC-2001 (TRENBLrel. 19, Created)		
DT	01-DEC-2001 (TRENBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TRENBLrel. 20, Last annotation: update)		
DE	A1562540/K19p1.15		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:		
OC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:		
OC	cucuridales 11: Brassicales: Brassicaceae: Arabidopsis.		
OC	NCBI_taxonomy:3702;		
RN	[...]		
RP	SEQUENCE FROM N.A.		
RA	Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,		
RA	Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,		
RA	Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,		
RA	Kamiya A., Karlin-Namand G., Kawai S., Lam H., Lee J.M., Lin J.,		
RA	Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Park C.J.,		
RA	Pham P.K., Quach H.L., Sakurai T., Satoh M., Seki M., Shinozaki K.,		
RA	Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,		
RA	Davis R.W., Theologis A., Ecker J.R.;		
RI	*Arabidopsis cDNA clones.*		
RI	Submitted (JUL-2001) to the EMBL/GenBank/DBS databases.		
CC	-1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER		
CC	PROTEINS (BY SIMILARITY).		
CC	-1- CATALYTIC ACTIVITY: AIP + UBIQUITIN + PROTEIN LYSINE - AMP +		
CC	DIPHOSPHATE + PROTEIN N-UBIQUITYLlysine.		
CC	-1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.		
CC	-1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN		
CC	THIOLESTER FORMATION (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.		
DR	EMBL: AY048267; AAK82529.1;		
DR	InterPro: IPR000508; URG_conjugat.		
DR	Pfam: PFC0179; UQ_con; 1.		
DR	ProDom: PD000451; URG_conjugat; 1.		
DR	PROSITE: PS00183; UBIQUITIN_CONJUGAL_1; CNKKNWN.1.		
DR	PROSITE: PS00127; UBIQUITIN_CONJUGAT_2; 1.		
DR	Ligase: Ubiquitin conjugation.		
DR	SEQUENCE 146 AA; 16625 MW; 1741F2A2C47045B3 CRC64;		
QY	7	ISSQALMLKLSQEEPEVEGRNITLVDSOLYNWEVAIFGLPNTLYEGSYFKAHKPP	56
DB	3	TPAKKRIMWLKRIQKDPVGIS-CAPCDNIHHNKLIFGSDPIPHCGGIFKLIHFTTE	61
QY	67	DYPSPTPEFELTKMHNPNYENGQVCTSIILHPDPOSGELPSERNNTQNRVAILLS	124
DB	62	DYNPKPTVREVSMEHPNYADGSGICLDL-----QNKSPIDYVAAVZS	108
QY	127	VSLSLNEPNTESPANDASVMFHKWRSKCKDKEYAFI	145
DB	109	IQSLCDPNPDSANAEARCFSE--NKKREYNKVFIEV	145
RESULT 40			
Q94SHC1			
ID	Q94SHC1	PRELIMINARY:	PRT: 152 AA.
AC	Q94SHC1		
DT	01-JUN-2002 (TRENBLrel. 21, Created)		

```

DT 01-JUN-2002 (TRENBLrei. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrei. 21, Last annotation update)
DR Rad6.
GN OSRA06.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCRL_TaxID-39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Yamamoto T., Shimanouchi K., Ishibashi T., Hashimoto J., Kimura S.,
RA Sakaguchi K.;
RT "Oryza sativa mRNA for Rad6, complete cds.";
RL Submitted (FEB-2002) to the FMBL/GenBank/DBJ databases.
DR EMBL; AB079798; BAB85469.1; -.
SQ SEQUENCE 152 AA; 17309 MW; ADAAA54032E237FD CRC64;

Query Match 17 88; Score 269; DB 10; Length 152;
Best Local Similarity 36.28; Pred. No. 1.4e-16;
Matches 59; Conservative 33; Mismatches 51; Indels 20; Gaps 4;

QY 7 TSSOKALMLELKSLOEPEVEGFRITLWDESDLYNNNEVAIFGLPNTLYEGGYFKAHIKPEI 66
   1 : : : | | : | : | : | : : : : | | : | : | : | : | : |
DB 3 TPARKRLMRDFKRLQODPPAGIS-GAPHNNIMLNNNAVIFGDDTPWDGGTFKLTLOFTE 61
   1 : : : | | : | : | : | : : : : | | : | : | : | : | : |

QY 67 DYPSPTFRFTKMHNNIYENGVCISILHPPVDDPQSGELPSRWNPNTONVRTILIS 126
   111 | | | : : : | : : | : : | : : | : : | : : | : : |
DB 62 DYPNKPVPVYRFVSRMFHFENIYADGSGICLDIL-----QNOWSPIDYVAAILTS 108
   111 | | | : : : | : : | : : | : : | : : | : : | : : |

QY 127 VISILNEPNTESPANVDASVMFRKWRDSKQKQKAEIIRKOV 169
   111 : | | : | | | : | : | : | : | : | : | : | : |
DB 109 IQLLCPNPNSPANSEARLF-----SENK-REYNRKVREIV 145
   111 : | | : | | | : | : | : | : | : | : | : | : |

```

Search completed: April 10, 2003, 10:36:15
 Job time : 49.7882 secs